Department of Psychology and Logopedics
Faculty of Medicine
University of Helsinki

WORKING MEMORY IN EXPLAINING INDIVIDUAL DIFFERENCES IN SCHOLASTIC SKILLS: INSIGHTS FROM ASSESSMENT AND TRAINING

Kaisa Kanerva

DOCTORAL DISSERTATION

To be presented for public discussion with the permission of the Faculty of Medicine of the University of Helsinki, in PIII, Porthania, on the 10th of October, 2019 at 12 o'clock.

Helsinki 2019

Supervisors

Dr. Virpi Kalakoski
Finnish Institute of Occupational Health
Helsinki
Finland

Professor Kimmo Alho
Department of Psychology and Logopedics
University of Helsinki
Finland

Reviewers

Docent Johanna K. Kaakinen
Turku Institute for Advanced Studies
Department of Psychology
University of Turku
Finland

Adjunct Professor Anna Soveri
Department of Clinical Medicine
University of Turku
Finland

Opponent

Professor Paola Palladino
Department of Brain and Behavioral Sciences
University of Pavia
Italy

ISBN 978-951-51-5422-4 (nid.)
ISBN 978-951-51-5423-1 (PDF)

Unigrafia
Helsinki 2019

ABSTRACT

Working memory (WM), a limited cognitive storing and processing mechanism for information, explains individual differences in scholastic learning and, more generally, complex cognition. One dominant view explaining these relationships proposes that variation in WM capacity reflects individual differences in general attentional processes. However, some models propose, that there are also domain-specific aspects in WM that explain variation in cognitive skills. Present series of four studies explored the nature of the mechanisms explaining the close relation between WM and scholastic skills in 5–16-year-old children and adolescents.

A total of 1069 subjects participated in the studies. Studies I and II investigated whether the individual differences in the general cognitive capacity are sensitive to external or internal distraction. In Study I the natural environmental noise at the classroom during WM assessment was documented (external distraction). In Study II the complexity of the WM task was manipulated (task's internal distraction). The contribution of the distraction in the memory scores and in the correlation between WM and scholastic skills was explored. Studies III and IV, in turn, aimed at investigating whether this relationship is causal, that is, whether the training of domain-general WM capacity or domain-specific counting skills, or both, would enhance children's WM and emergent mathematical skills. In two interventions lasting four to five weeks, domains of WM components (verbal WM and short-term memory, STM; visuospatial WM and STM; Study III) and domains of outcome (counting, combined WM and counting; Study IV) were addressed.

The results of Studies I and II showed that environmental distraction and task demands contributed the relationship between WM and scholastic skills manifesting the individual differences more clearly. The cognitive constructs assessed appeared to be highly overlapping. However, the results of Studies III and IV indicate that while WM has an important role in scholastic skills, the computerised training of different WM domains did not lead to improvement in numeracy. Despite the lack of such training effects, the group-based interventions addressed to the skill of interest, in this case numeracy, enhanced these skills. Taken together, the present results suggest that while attentional load contributes substantially to individual differences in WM capacity by restricting the mental workspace, the acquired long-term memory representations are needed in order to apply the WM capacity in scholastic learning. The results of the present thesis can be applied in recognising the cognitive deficits that hinders childrens' scholastic learning, and in developing interventions that benefit scholastic skills.

TIIVISTELMÄ

Työmuisti on kapasiteetiltaan ja kestoltaan rajallinen tiedon säilytys- ja prosessointimekanismi. Yksilöllinen vaihtelu työmuistin toiminnassa selittää yksilöiden välisiä eroja kouluoppimisessa ja muissa vaativissa kognitiivisissa eli tiedonkäsittelyn tehtävissä. Yhden vallalla olevan selitysmallin mukaan työmuistin kapasiteetin yksilöllinen vaihtelu heijastaa yksilöiden välisiä eroja yleisissä tarkkaavuuden prosesseissa. Toisten mallien mukaan työmuistiin kuuluu erikoistuneita alayksiköitä ja tiedonkäsittelyn toimintoja, jotka selittävät yksilöiden välisiä eroja kognitiivisissa taidoissa. Tämän väitöskirjatyön neljän osatutkimuksen tavoitteena oli selvittää tarkemmin selittävätkö yleiset vai erikoistuneet työmuistitoiminnot työmuistikapasiteetin ja koulutaitojen yksilöllisen vaihtelun vahvaa yhteyttä 5–16-vuotiailla lapsilla ja nuorilla.

Kaikkiaan 1069 koehenkilöä osallistui tutkimuksiin. Osatutkimukset I ja II selvittivät, ovatko yksilölliset erot kognitiivisessa kapasiteetissa herkkiä ulkoiselle tai sisäiselle häiriölle. Ensimmäisessä osatutkimuksessa kirjattiin ulkoisen ympäristön eli luokkahuoneen melu työmuistiarvion aikana (ulkoinen häiriö). Toisessa osatutkimuksessa käytettiin työmuistitehtäviä, joissa oli vaativuustasoltaan erilaiset prosessointiosiot (tehtävän sisäinen häiriö). Ulkoisen ja sisäisen häiriön vaikutuksia tutkittiin toisaalta suhteessa työmuistitehtävässä suoriutumiseen ja toisaalta suhteessa työmuistikapasiteetin ja koulutaitojen välisen yhteyden voimakkuuteen. Osatutkimukset III ja IV sen sijaan pyrkivät selvittämään, onko työmuistikapasiteetin ja koulutaitojen suhde kausaalinen, eli parantaako yleisen työmuistikapasiteetin ja/tai erityisen laskemistaidon harjoitteleminen lasten työmuitikapasiteettia ja varhaisia matemaattisia taitoja. Vaikuttavuustutkimuksissa harjoitettiin tietokoneavusteisesti viiden viikon ajan työmuistin erikoistuneita osa-alueita (kielellinen työmuisti ja lyhytkestoinen muisti; visuospatiaalinen työmuisti ja lyhytkestoinen muisti; Osatutkimus III) sekä ryhmämuotoisesti neljän viikon ajan harjoitettavan taidon kannalta oleellisia taitoja (laskeminen; sekä työmuisti että laskeminen; Osatutkimus IV).

Osatutkimusten I ja II tulokset osoittivat, että sekä ympäristön häiriöt että tehtävän vaatimukset vaikuttivat työmuistikapasiteetin ja koulutaitojen välisen suhteen voimakkuuteen siten, että yksilöiden väliset erot tulivat esiin tietyissä ulkoisen ja sisäisen häiriön tilanteissa selvemmin kuin tilanteissa, joissa häiriö oli vähäisempi. Lisäksi arvioitujen kognitiivisten toimintojen havaittiin olevan päällekkäisiä. Osatutkimukset III ja IV puolestaan osoittivat, että vaikka työmuistilla on vahva yhteys koulutaitoihin, työmuistin eri alayksiköiden harjoittaminen ei vahvistanut lasten varhaisia matemaattisia taitoja. Sen sijaan laskemisen harjoittaminen vahvisti lasten varhaisia matemaattisia taitoja. Tämän väitöskirjan tulosten kokonaisuudesta voidaan

päätellä, että työmuistin yleisen tarkkaavuusmekanismin kuormittuminen kaventaa mentaalista työtilaa ja on siten merkittävä tekijä yksilöiden välisen kognitiivisen suoriutumisen selittäjänä. Kuitenkin myös aiemmin opittuja tehtäväkohtaisia säilömuistin edustuksia tarvitaan, jotta työmuisti voi toimia tehokkaasti ja edistää näin kouluoppimista. Tutkimustuloksia voidaan hyödyntää lasten kouluoppimisen taustalla olevien ongelmien tunnistamisessa sekä koulutaitoja edistävien harjoitteiden kehittämisessä.

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my supervisors, Dr. Virpi Kalakoski and Professor Kimmo Alho. Virpi has not only been vital for accomplishing the current thesis, but has mentored me throughout literally all my time studying cognitive psychology and research methods. By offering encouragement, support and independence in the perfect proportions she has provided me the most rewarding journey in scientific thinking that one could wish for. I am also grateful to Kimmo for constructive, thorough guidance that has integrated me into the scientific community. I also thank Kimmo for taking on the role of Custos. I am grateful to Professor Paola Palladino for agreeing to be my Opponent and to Docent Johanna Kaakinen and Docent Anna Soveri for reviewing my thesis, giving new insights for my work.

I wish to thank all my co-authors, with whom this work has been not only compelling, but also a great learning opportunity. First, I wish to thank Docent Minna Kyttälä in the University of Turku. I am grateful for Minna for giving me a huge opportunity to learn by sharing the projects with me and for prompt, efficient support for my work whenever I needed it. I also wish to thank our co-author, Docent Evelyn Kroesbergen from the Utrecht University. I wish to thank Docent Noona Kiuru and Dr. Riikka Hirvonen from the University of Jyväskylä, and their whole team together with Emeritus Professor Timo Ahonen for providing me the possibility to be part of their enthusiastic research group, for smooth and positive collaboration, and for always being ready to help and support my work.

I am deeply grateful for all the valuable help I have received during this work from the supportive and technical staff and students at the University of Helsinki, including explaining statistics, writing computer code for experiments and interventions, collecting and managing data and commenting on manuscripts. Special thanks go to Pekka Lahti-Nuuttila, who has always been a great support, not only for my research, but also in other big efforts that we have been involved in. I also wish to thank Marja-Leena Haavisto for the beneficial and insightful discussions and feedback for my work. I thank Milla Jokinen for proofreading this thesis.

With warm gratitude I wish to thank all the children and adolescents for participating in the research, and teachers and schools for providing time and space for conducting this research. Financial support from the University of Helsinki, Otto A. Malm Foundation, Emil Aaltonen Foundation, Finnish Brain Foundation, Finnish Cultural Foundation and Academy of Finland has provided me and the research groups with the opportunity to conduct this research, which is thankfully acknowledged.

At the Department of Psychology and Logopedics, in its various forms during this work, there are, or have been researchers, who deserve special thanks. First, I wish to thank Docent Elisabet Service for the insightful discussions, feedback and support during whole this project. Second, I wish to thank Docent Markku Verkasalo for the most inspiring atmosphere he creates wherever he is. During this work his example has encouraged me to find my own interests and to always follow them. He has also provided me with the opportunities of, at times much needed, paid work. Third, I am happy to have a possibility to thank Docent Jussi Saarinen for his warm support and encouragement during my work, especially during the beginning of my academic career. I also thank Jussi for taking the role of the faculty representative for the public examination of my thesis.

I wish to thank all the reviewers of my papers, all co-authors in papers not presented here, participants of the Everything is Fucked -reading club and the good old WM seminar. During my postgraduate studies I have learned the most from you.

I am lucky to have had a possibility to combine this scientific work with the more real and practical life of children as a clinical (neuro)psychologist. I wish to thank all my colleagues and friends in various places I have worked at as a psychologist, and all the fellow test-enthusiastics at Hogrefe Psykologien Kustannus.

While in the scientific collaboration there is sometimes a need to prove yourself, this is never the case with friends and family. I am lucky to have such inspiring friends outside and inside academia. Special thanks are reserved for Petra Nyman, who was there when this all started. Time flew during our discussions trying to understand the questions at hand not ending before we both had shared understanding of it. Next, I wish to thank Jenni Heikkilä, who in turn has shared the combination of research and clinical work with me. With Jenni I have been able to share not only the trickiness and delights of our work combo but also career plans and new ideas that it produces. I am thankful to my friends and relatives, who have helped me to change my focus from research on practical things, like my cousin Hanna at our mökkiviikko. I am grateful to my goddaughter Nora for making the last bits of this work possible by staying at our place and taking care of our smallest one. I am thankful to Milla and Aara for being there for their little sister.

Äiti ja isä, thank you for always letting me make my own decisions. This is what it led to.

Finally, my dearest ones at home. Ile, I should thank you for the collaboration in this project, but what I really am thankful for is that you have built a home and a life with me, always having new and bright insights for everything that we do together, opening up my eyes time after time. I wish to thank Eeris, who grew up together with this thesis, and Inna Myy, who grows up independently of this thesis, or anything, it seems. I dedicate this thesis to you, Eeris and Inna Myy. I hope that it will give you the passion and courage for discovering - whatever it could be, we do not yet know.

Helsinki, 1st September, 2019
Kaisa Kanerva

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I Kanerva, K., Kiistala, I., Kalakoski, V., Hirvonen, R., Ahonen, T., & Kiuru, N. (2019). The feasibility of WM tablet tasks in predicting scholastic skills in classroom settings. *Applied Cognitive Psychology*. doi:10.1002/acp.3569.

II Kanerva, K., & Kalakoski, V. (2016). The predictive utility of a working memory span task depends on processing demand and the cognitive task. *Applied Cognitive Psychology*, 30, 681-690. doi:10.1002/acp.3243.

III Kanerva, K., & Kyttälä, M. (2016). Specific training of working memory and counting skills in kindergarten. *Cursiv, 18,* 159-176.

IV Kyttälä, M., Kanerva, K., & Kroesbergen, E. (2015). Training counting skills and working memory in preschool. *Scandinavian Journal of Psychology*, 56, 363-370. doi:10.1111/sjop.12221.

The publications are referred to in the text by their roman numerals.

1 INTRODUCTION

Working memory (WM) is a limited-capacity cognitive storage and processing mechanism for information needed temporarily available when we are engaged by an ongoing thought or action (Cowan, 2014; 2017). The basic elements of WM are maintenance of information available for further processing, its temporary nature, and its limited capacity. Thus, WM has been described as a temporary storage and processing system which functions under attentional control and which underpins our capacity for complex thought (Baddeley, 2007).

There are two main reasons why WM is important in academic learning. First, the capacity of WM is restricted, which affects our ability to retain, select and integrate novel information together with previously acquired information in the long-term memory (LTM) (see Cowan, 2014). For example, when calculating verbal mathematical problems, one must retain the numbers and the required operation in memory in order to perform the task. However, if task demands exceed the WM capacity, some information is lost and the task will not be completed correctly. In other words, the amount of WM capacity required by a learning task affects the level of task performance and assessed learning outcome.

Second, WM is important in academic learning, because of the individual differences in the capacity to hold information in WM (Cowan, 2014; Engle & Kane, 2003). Scholastic learning, for example learning to read and learning mathematical skills, relies on WM, and with lower capacity learning outcomes are impaired (Cowan, 2014). Several studies demonstrate that individual differences in WM capacity predict a wide variety of complex cognitive tasks, such as reasoning (Engle, 2018; Gignac, 2014; Kyllonen & Christal, 1990), but also school relevant skills (Alloway, 2006; Fenesi, Sana, Kim, & Shore, 2015; Peng et al., 2018), such as second language learning (Palladino & Ferrari, 2008) and perspective taking during reading (Kaakinen, Hyönä, & Keenan, 2003). These associations are documented both in children and in adults. Thus, impaired performance in a cognitive task may reflect individual differences in holding information in WM rather than problems with other cognitive functions or skills.

According to previous research, it is clear that WM capacity is related to scholastic skills in children and adolescents of different ages. However, there are two important open questions regarding the role of WM in scholastic skills. The first question is, why the measured WM span explains cognitive skills, such as scholastic skills so well. One possible explanation is related to the controlled attention needed in both WM and scholastic skills (Engle & Kane, 2003). In the present thesis, the attentional requirements of WM processing are addressed by applying external (Study I) and internal (Study II) distraction in order to evaluate whether distraction would affect the correlations between the measured WM span and scholastic skill. If executive attentional processes are involved, internal or external distraction, or both, should affect the span scores or the correlations between WM and scholastic skills (i.e., the predictive ability of WM task), or both.

The second open question regards the domain-specificity of the relationship between WM and scholastic skills. The domain-general view suggests that since WM is assumed to serve as a general capacity supporting scholastic skills, WM training would improve those skills (Melby-Lervåg & Hulme, 2013). The domain-specific view, in turn, suggests that WM is related to scholastic skills because of shared knowledge, and the knowledge base has to be strengthened in order to improve particular scholastic skill (Ericson & Kintsch, 1995). The high correlation between the measured WM span and scholastic skills obtained in previous studies proposes that by training WM, scholastic skills, such as mathematical skills, should improve. However, the mechanisms of transfer are less known. The present thesis addresses two possible venues that could mediate the effects of interventions of scholastic skills: the domain of the trained WM component (Study III) and the domain of the outcome (Study IV). If WM is purely a domain-general capacity supporting learning, training WM in general or particular components of WM should enhance mathematical skills. If, however domain-specific aspects are crucial in explaining scholastic skills, effects would manifest when training of the domain of outcome.

Since WM capacity has been shown to explain scholastic skills throughout developmental stages (Gathercole, Pickering, Knight, & Stegmann, 2004, Nevo & Bar-Kochva, 2015), these effects are assumed to be evident despite of a child's age. Thus, the relationship between WM capacity and scholastic skills or emergent scholastic skills, should be present at kindergarten and throughout primary and secondary education.

1.1 WM IN SCHOOL LEARNING

Variety of definitions and models have been proposed during the over 40 years of modeling the architecture of WM and its role in everyday cognition (for reviews, see Cowan, 2014; Conway, Jarrold, Kane, Miyake, & Towse 2007; Miyake & Shah, 1999). There is a general agreement among researchers that WM consists of domain-general and domain-specific functions (Baddeley & Hitch, 1974; Engle, 2002). This means that on one hand, irrespective of the nature of the material, there is a certain general executive capacity in use. However, on the other hand, some WM functions are specific to the nature of the material, that is, whether it is verbal or visuospatial, and some functions are specific to content knowledge related to the task. What is unsettled among researchers is what is the role of WM in explaining the relationship between WM and complex cognitive skills, such as scholastic skills. That is, whether purely the domain-general component or a combination of domain-general and domain-specific components are responsible for this relationship.

1.1.1 DOMAIN-GENERAL VIEW OF WM EXPLAINING SCHOLASTIC SKILLS

One dominant view explaining the relationship between WM and complex skills, executive attention account, proposes that variation in WM capacity reflects individual differences in general attentional processes, that is, the ability to maintain and process information in distracting settings (Engle, 2002; 2018; Kane, Conway, Hambrick, & Engle, 2007). According to this view, the WM capacity reflects the ability to access relevant information, when interference is high (see Shipstead & Engle, 2013; Unsworth & Engle, 2007). The conclusions of the executive attention account are based on research combining experimental and correlational approaches and studies that have used quasi-experimental approaches to compare individuals with high and low span (Kane et al., 2007). For example, in their seminal studies, Engle and his colleagues (Kane, Bleckley, Conway, & Engle, 2001) first identified the individuals who were in the upper quartiles (high spans) and lower quartiles (low spans) on WM capacity measured by complex span tasks, and subsequently assess them with other tasks to see whether they perform differently in various cognitive tasks. Studies have demonstrated that high and low-span individuals differ in performing large variation of real-world tasks, including reading and listening comprehension, vocabulary learning and writing (see Engle, 2001). However, WM capacity does not predict performance in a task, in which participants are required to make a saccade toward (prosaccade), in contrast to making it away (antisaccade) from a flashing cue (Unsworth, Schrock, & Engle, 2004). Thus, it has been suggested that controlled, rather than automatic attentional processes explain individual differences. More recently, latent-factorial methods have been adopted to support the conclusions of the role of individual differences in controlled attention (see Kane et al., 2007).

In scholastic learning, in which several kinds of distractors are present in the environment and in the tasks themselves, understanding their influences is practically and theoretically relevant. The executive attention account proposes that distraction would disrupt WM performance since WM is responsible for inhibiting irrelevant stimuli. This would be expected for both external and internal distraction, due to their demands for controlled attention and inhibition control. Therefore, Studies I and II of the present thesis address the questions of the contribution of external and internal distraction to the relationship between the measured WM span and scholastic skills in both laboratory and natural classroom settings.

1.1.2 WM AS A DOMAIN-SPECIFIC MECHANISM

While WM is seen as a domain-general capacity, some models have proposed that there are also domain-specific aspects in WM that explain the variation in scholastic skills between individuals or specific groups, for example, children with various learning difficulties. First, the multicomponent model (Baddeley, 2002; Baddeley & Hitch, 1974) suggests that WM consists of two separate, domain-specific short-term memory systems; the phonological loop and the visuospatial sketchpad. They are short-term in a sense that they are responsible for temporary maintenance of information. Domain-general components of WM, that is, the central executive and the episodic buffer, in turn, coordinate ongoing processing and storage of verbal and visuospatial information, and bind novel verbal and visual representations together and to those stored in the LTM (Baddeley, 2002).

A wide literature exists in evaluating the roles of WM components in various scholastic skills, including mathematics (Imbo & Vandierendonck, 2008; Raghubar, Barnes, & Hecht, 2010), reading (Swanson, Howard, & Saez, 2006) and second language learning (Linck, Osthus, Koeth, & Bunting, 2014; Palladino & Ferrari, 2008; Service, 1992). For example, Logie, Gilhooly and Wynn (1994) revealed with dual-task methodology that articulatory suppression increased errors in mental arithmetic, showing the essential role of phonological loop in mental arithmetic.

The other model emphasizing the role of domain-specific information stresses that the acquired domain knowledge in LTM mediate the relationship between WM and complex skills (Ericsson & Delaney, 1999; Ericsson & Kintsch, 1995). According to Ericsson and Kintsch (1995), LTM can facilitate and support WM processing. The model assumes that when individuals have prior knowledge in a particular domain, they can encode and retrieve information specific to it more efficiently than when they have less prior information in a given domain. This theory proposes that variation in WM is due to differences in the ability to encode task-specific information into LTM and use retrieval cues to rapidly access task-specific knowledge in LTM (Ericsson & Delaney, 1999; Ericsson & Kintsch, 1995).

In sum, there is variation in WM models, and the attentional aspects of domain-general system are emphasized in the majority of current models. However, the domain-specific prior knowledge on the other hand is critical in scholastic learning and thus its role in explaining the relationship betweeen WM and scholastic skills is interesting. The current thesis addresses in the context of scholastic learning in children and adolescents, first, the attentional mechanisms of domain-general WM (Studies I and II), and second, the domain-specific aspects of WM (Studies III and IV).

1.1.3 ASSESSMENT OF WM WHEN PREDICTING SCHOLASTIC SKILLS

WM is typically assessed with span tasks, which basically require encoding an increasing number of presented items and in the end, retrieving the encoded material. The measured WM span reflects thus the highest amount of information that the participant is able to reliably retrieve, and more items recalled in span tasks suggest higher WM capacity. The assessments are traditionally conducted individually in controlled laboratory settings to avoid measurement error due to the variations in testing situations.

The convention in the field is to refer storage-oriented span tasks with no explicit concurrent processing as simple short-term memory (STM) span tasks, and span tasks that involve explicit concurrent processing requirement together with storage as complex WM span task (Bayliss, Jarrold, Gunn, & Baddeley, 2003; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). Thus, the two most commonly adopted span tasks are 1) the simple span tasks, which aim at assessing more pure maintenance capacities (storage and rehearsal/refreshing) (Alloway, 2007) and 2) the complex span tasks, which aim at assessing concurrent processing and storage functions (Conway et al., 2005). Some researchers have proposed that since the simple span tasks assess pure storage functions, they should be considered STM measures. The complex span tasks assess not only storage functions, but additional processing functions and should thus be considered WM measure. In present thesis, WM is considered as processing and storage system, in which pure storage functions are called STM.

Furthermore, both simple and complex span tasks can be devoted in assessing the maintenance of verbal or visuospatial information. A set of verbal and visuospatial simple and complex span tasks have been collected as comprehensive standardised test batteries, namely, Automated Working Memory Assessment (AWMA; Alloway, 2007) and Working Memory Test Battery (WMTB-C; Pickering & Gathercole, 2001), which both are adopted in the current Studies III and IV to obtain a broad approximation of childrens WM. These batteries are designed for clinical, educational and research purposes, and consists of several widely used verbal and visuospatial span tasks.

Complex span tasks are currently widely used and most researched WM tasks and they are considered as a gold standard of WM measurement (Towse & Hitch, 2007; Cowan, 2017). The complex span tasks (also called the WM span tasks) are a set of different tasks in which the participant needs to store items in memory while conducting a processing task in between the to-be-memorized items (Daneman & Carpenter, 1980; Turner & Engle, 1989). Various memory and processing components have been adopted in studies since Daneman and Carpenter (1980) introduced their first Reading Span task.

In the original Reading Span task, reading sentences was used as a processing task, and retaining the last word of each sentence as a memory component. Other widely used processing components have been tasks of counting dots (Counting Span; Case, Kurland, & Goldberg, 1982) and verifying arithmetic operations (Operation Span; Turner & Engle, 1989). Commonly used memory components, in turn, have involved words (Daneman & Carpenter, 1980), digits (Turner & Engle, 1989) or letters (Barrouillet, Bernardin, & Camos, 2004) unrelated to the processing task. Performance in different processing components of complex WM tasks have been shown to systematically predict higher-order cognition to a fairly similar degree. In studies in children and adolescents, two commonly adopted complex span tasks are the Counting Span task (Case et al., 1982) and Reading Span task (Geers, Pisoni, & Brenner, 2013). The complex span tasks have been demonstrated to be highly reliable, indicated by high test-retest correlations (Klein & Fiss, 1999) and to have a high construct validity (Conway et al., 2005).

The close association with scholastic skills, and the sensitivity to attentional demands shown in previous studies (Bunting, 2006; Magimajraj & Montgomery, 2010; St. Clair-Thompson, 2007) were relevant aspects when selecting complex span tasks as assessment tools in Studies I and II of the present thesis. First of all, the potential of complex WM span tasks in assessing WM in school context is supported by results showing that complex span tasks tend to be stronger predictors of cognitive activities than simple span tasks (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999). Simple span tasks account for no unique variance in general cognitive ability after variance related to the measured WM span is accounted for (Engle, Kane, & Tuholsky, 1999). Further, when presented with a computer, the presentation pace and content of the complex span tasks processing component can be manipulated in order to investigate the factors contributing to the predictive utility of the task (McCabe, 2010).

1.1.4 WM AND LEARNING AT SCHOOL: CORRELATIONAL EVIDENCE

WM capacity is related to scholastic skills throughout the childhood development, starting with emerging academic skills, such as emergent literacy and numeracy, at preschool age (Alloway et al., 2005; Preßler, Krajewski, & Hasselhorn, 2013) and continuing in learning at secondary school (Gathercole et al., 2004). Furthermore, longitudinal studies have demonstrated that WM assessed in early childhood predicts later academic skills (Alloway & Alloway, 2010; Kyttälä, Kanerva, Münter, & Björn, 2019). The importance of WM in school learning is further emphasized by showing WM deficits in learning difficulties (Gathercole et al., 2016; Swanson & Sachse-Lee, 2001), such as developmental dyslexia (Kudo, Lussier, & Swanson, 2015; Menghini, Finzi, Carlesimo, & Vicari, 2011) and mathematical learning difficulties/dyscalculia (Peng & Fuchs, 2016).

Most important aspects of scholastic learning in which WM has documented to play a role as a domain-general capacity are mathematical skills (Swanson, 2011), reading comprehension (Daneman & Carpenter, 1980), performance in the national curriculum tasks and tests (Gathercole et al., 2004) and the grade point average (GPA; Gathercole, Brown, & Pickering, 2003; Gathercole & Pickering, 2000). To describe the role of WM in mathematical skills, Peng and his colleagues (Peng, Namkung, Barnes, & Sun, 2016) explored the correlations between the WM capacity and mathematical skills in a recent meta-analysis of 110 studies. The results revealed that the measured WM span correlated positively with mathematical skills across the studies (r = .35). Furthermore, while the type of mathematical task affected the strenghth of the relation between WM capacity and mathematics, no moderating effect of domain of WM span task was found. That is, the relation was domain-general, since verbal, numerical, and visuospatial WM spans predicted similar amount of variance in mathematics.

WM capacity is a strong predictor not only of scholastic skills, but also of fluid intelligence (Ackerman, Beier, & Boyle, 2002; 2005; Conway et al., 2002; Conway & Kovacs, 2013; Kyllonen & Christal, 1990; Süss, Oberauer, Wittmann, Wilhelm, & Schulze, 2002). Fluid intelligence is defined as reasoning with novel information, as opposed to reasoning with domain-specific, previously acquired knowledge (McGrew, 2009). However, also fluid intelligence is related with scholastic skills. Several studies show that fluid intelligence and scholastic performance correlate (e.g., Deary, Strand, Smith, & Fernandes, 2007; Krumm, Ziegler, & Buehner, 2008; Rohde & Thompson, 2007; Vock, Preckel, & Holling, 2011). In a recent meta-analysis of 680 studies, fluid intelligence correlated moderately with both mathematics and reading, the relations being stronger to mathematics than to reading (Peng, Wang, Wang, & Lin, 2019). Furthermore, fluid intelligence and academic skills significantly predicted each other in the development, which suggest that the relationship between fluid intelligence and academic skills is reciprocal (see also Kievit et al., 2017).

Despite of the close relationship between fluid intelligence and scholastic skills, WM uniquely contributes to explaining scholastic achievement after fluid intelligence is accounted for (Alloway & Alloway, 2010; Lu, Weber, Spinath, & Shi, 2011; Maehler & Schuchardt, 2009; Swanson, Jerman, & Zheng, 2008). Recent research in adults have tried to separate the roles of WM and fluid intelligence in higher order cognition, which is difficult with the highly correlated constructs (Engle, 2018). Work on this area has demonstrated that in adults WM capacity is related to fluid intelligence assessed with non-verbal intelligence tests independent of processing speed (Redick et al., 2012).

Taken together, because of its close relationship between WM and scholastic skills, fluid intelligence is an important variable to take into account and to extract from WM, when exploring the relationships between WM and scholastic skills. Moreover, it would be important to be able to describe the common and unique contributions of WM and fluid intelligence to scholastic learning. In Studies I and II of the present thesis, the associations between

WM, fluid intelligence and scholastic skills are explored in adolescents by adopting complex span tasks for assessing WM.

1.2 FACTORS CONTRIBUTING TO THE PREDICTIVE UTILITY OF WM SPAN TASKS

In order to understand why WM capacity, especially when assessed with complex span tasks is such a good predictor of complex cognition, the framework of studying the factors affecting the measured WM span and the tasks predictive utility have been adopted. There are two basic approaches: The first is to manipulate or take into account the external factors, such as distracting noise in the environment and evaluate its contribution to the measured WM span (Elliott, 2002). This approach was adopted in the Study I. The second is to manipulate the task itself, that is, the task's internal properties, such as the demand of the processing component (Lépine et al., 2005). This approach was adopted in the Study II. The outcome of interest has typically been the WM span score (Barrouillet et al., 2004; Elliott et al., 2016), but even more interesting outcome is the correlation between the WM span score and scholastic skills. The idea behind this approach is that if the change in the circumstances of WM assessment is accompanied by a change in the WM span score or in the correlation between the span score and scholastic skills, it can be interpreted as reflecting critical aspects of individual differences in WM capacity or scholastic skills.

1.2.1 EXTERNAL FACTORS CONTRIBUTING TO THE RELATIONSHIPS BETWEEN WM AND SCHOLASTIC SKILLS

Ambient distractors, such as environmental noise during WM processing may affect the performance in a WM task or the task’s ability to tap the underlying cognitive construct (Redick et al., 2012), and most importantly such distractors may also contribute to the relationship between WM and scholastic skills. Previous research has demonstrated that even brief exposure to an irrelevant sound can be detrimental to the cognitive performance of children and adults (Sörqvist, 2010; Vasilev, Kirkby, & Angele, 2018). Detrimental effects of irrelevant auditory distractors, verbal or non-verbal, on performance in WM tasks are referred to as irrelevant sound effects (Salamé & Baddeley, 1982). Any extraneous stimulation may constitute a potential source of distraction (Wetzel & Schröger, 2007). For example, both verbal and non-verbal auditory distraction has been documented to disrupt memory performance (Tremblay, Nicholls, Alford, & Jones, 2000), also in classroom settings (Hygge, 2003). However, some research indicates that irrelevant speech distracts more than other irrelevant sounds (Salamé & Baddeley, 1982). Furthermore, children are found to be more susceptible to disruptive effects of irrelevant sounds than adults (Elliott, 2002; Elliott & Briganti, 2012; Elliott et al., 2016).

The disruptive nature of environmental noise on performance in children and adults has been explained by the requirements of attentional control (Engle, 2001). This ability improves in development (Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2010), which could explain the adult's relative superiority in inhibiting the effect. In the context of the assessment of WM in a school classroom, an important question, investigated in Study I is, whether the environmental noise affects the reliability of WM assessment or the measured relationship between performance in WM tasks and other cognitive tasks (e.g., scholastic tasks or fluid intelligence tasks).

1.2.2 TASK'S PROPERTIES CONTRIBUTING TO MEASURED RELATIONSHIPS BETWEEN WM AND SCHOLASTIC SKILLS

To assess the contribution of internal properties of WM span tasks to their predictive utility, the approach of evaluating the effect of processing component of complex WM tasks on participants' item recall has been adopted. The basic idea is that manipulating the aspects of the WM span task's processing component, rather than those of the storage component, affect the measured WM span and determine the predictive value of this span (Barrouillet et al., 2004; Lépine et al., 2005). The processing component aspects have been studied in varying its cognitive demands, for example, the difficulty of math operations (Bunting, 2006; Turner & Engle, 1989) when using the operations as processing tasks, or the semantic complexity of sentences (Magimairaj & Montgomery, 2012) when using sentences as processing tasks.

Some studies suggest that the attentional demands of the WM processing task, not the processing time or the task's difficulty per se, determine the predictive utility of a WM task. However, studies comparing different WM processing tasks in predicting higher order cognition have produced conflicting results depending on the nature of the complex span task (i.e., Bunting, 2006; St. Clair-Thompson, 2007). Furthermore, it is possible that the ability being predicted has some role in the relationship. These questions are addressed in Study II.

In sum, it is widely accepted that WM has a strong role in complex cognition. The potential contributors to the relationship between the measured WM span and scholastic skills are environmental noise and the processing task demands, which are both related to attentional requirements of task administration. However, it is not clear whether the effects of environmental noise or processing task demands are positive or negative and whether they are robust across the cognitive skills to be predicted. These questions were addressed in the present thesis in Studies I (external distraction; i.e. environmental noise) and II (internal distraction, i.e.

processing task demands) in the context of school learning in 12- and 15-year-old adolescents.

1.3 IS THERE A CAUSAL RELATIONSHIP?

The correlational relationship between WM performance and other cognitive performance has been well documented, and also in longitudinal studies WM performance has been documented to predict later academic achievement (Kyttälä et al., 2019). The tight correlational relationship between WM performance and other cognitive performance raises the question of whether improving WM with training could increase the domain-general attentional capacity, and would this be manifested in improved performance in other tasks in which domain-general attentional capacity is critical (Melby-Lervåg & Hulme, 2013).

Although the correlational relationship between WM performance and cognitive skills is clear, cross-sectional studies showing this association do not confirm a causal relationship between WM and cognitive activities (Sala & Gobet, 2018). One way for establishing a causal relationship is to show that WM training leads to improvement in scholastic skills. If by enhancing WM capacity we can improve scholastic skills when other aspects are carefully controlled, the enhanced WM capacity can be interpreted to be the cause of improved scholastic skills. Naturally, the other purpose of WM training studies has been much more practical, reflecting the hope for enhancing learning potential and support children and adults struggling with limited capacity of WM in their learning and in everyday life. In present thesis however, a more theoretical viewpoint on this highly debated issue is adopted.

1.4 WORKING MEMORY TRAINING AND TRANSFER

The standard assumption underlying training studies states that WM serves as domain-general ability that may be enhanced by training, and that these improved cognitive abilities transfer to other domains. Typically, training involves repetition of demanding WM task, or a set of WM tasks for a specific amount per session for a specific number of scheduled sessions (e.g., Klingberg, Forssberg, & Westerberg, 2002). Furthermore, the training paradigms are commonly designed to demand high cognitive workloads and thus adapt to participants' varying level of proficiency. During a decade of investigations, an enormous number of studies has been published on WM training, and on brain training more generally, trying to reveal training effects on other cognitive skills in non-clinical and clinical populations (for meta-analysis, see Au et al., 2015; Melby-Lervåg & Hulme, 2013; Schwaighofer, Fischer, & Bühner, 2015, Soveri, Antfolk, Karlsson, Salo, & Laine, 2017).

1.4.1 NEAR AND FAR TRANSFER

The goal of the WM training is to elicit positive transfer on untrained tasks. Thus, the interest has been on the existence of generalization following cognitive training. Transfer refers to the effect that the practicing of one task has on the performance of another task (Barnett & Ceci, 2002). The transfer-effects are considered as near-transfer, when the transfer is seen between two structurally similar tasks, that is, practicing with a verbal span task affects performance in the same verbal span task even if the memory items in the task are replaced with new items. In contrast, the effect is considered as far-transfer, if the transfer is seen between two structurally different tasks, that is, when practicing verbal word span task affects performance in verbal counting of numbers task.

The experimental evidence regarding the influence of WM training on cognitive function has been inconsistent. While some studies have suggested that WM training programs enhance particular aspects of cognition (Au et al., 2015), others have claimed the opposite (Melby-Lervåg & Hulme, 2013). The meta-analyses conducted in the field mostly agree that WM training effects generalise to tasks that are similar to the trained tasks, but not to tasks that are different (Melby-Lervåg & Hulme, 2013). Other meta-analyses in healthy adults (Soveri et al., 2017), typically developing children and young adolescents (Sala & Gobet, 2017a) support this conclusion. Thus, while near-transfer has been commonly shown, research is more cautious on whether far-transfer can be obtained with WM training.

1.4.2 PROPOSED MECHANISMS FOR TRANSFER

Most of the studies to date have been devoted to answering merely the question of whether, with a specific training program in a specific target group, there are training effects or not. Few studies have focused on revealing the mechanisms of WM training (Gathercole, Dunning, Norres, & Holmes, 2019; von Bastian & Oberauer, 2013). On the neural level the proposed mechanisms are related to plasticity of the brain networks (Klingberg, 2010; Takeuchi, Taki, & Kawashima, 2010). On the level of cognitive processes, it has been suggested that the training-induced broad cognitive improvements can be caused by either increased WM capacity or enhanced WM efficiency, or a combination of both (von Bastian & Oberauer, 2014). However, the domain-general and domain-specific aspects of training are not clear. In order to evaluate the mechanisms of WM training, the proposed mechanisms should be identified and addressed in separate experimental groups when studying the training effects.

In present thesis, the domain-specific aspects of training were addressed in two studies, from two different perspectives: (1) the domain and the complexity of the training tasks (verbal STM/WM, or visuospatial STM/WM), and (2) the domain of outcome (counting or combined WM and counting). Consequently, first approach, adopted in Study III, was to address domains of the WM separately by training each component in separate group and evaluating the effects of training on 1) corresponding WM components, 2) across WM components and 3) transfer to more complex skill, in this case numeracy. The second approach, adopted in Study IV, was to compare the effects of training of WM with the effects of training a domain that is similar to the domain of interest. More specifically, as the goal was to improve children's early numeracy, the training was addressed to WM in one group and to early counting skills in the other group.

1.4.3 CONCERNS IN TRAINING STUDIES

Many concerns have been raised regarding not only aforementioned theoretical but also methodological aspects of WM training studies. One methodological issue that has raised concerns in a large body of studies of WM training is the active vs. passive nature of the control group (von Bastian & Oberauer, 2013), that is, are the participants in the control group performing any tasks during the intervention period (active control group) or is the control group only participating in the pre- and post-assessments (passive control group). Both active and passive control group control for possible retest effects. That is, using an untrained group as a comparison group allows evaluating whether the experimental group improves more than the control group in particular skills. However, active controls are needed to control for other generic intervention effects, such as being a participant in an intervention, using computer or working with the experimenter (von Bastian & Oberauer, 2014). These issues could be particularly important when studying small children, who have less experience of testing situations. Furthermore, in order to control for effects of expectations and motivation, the control activity of the active control group should be believable and the support and feedback in the control group’s activity should correspond to those in the training task, so that the control activity and training task would not differ from each other in their motivational aspects. In Study III of the present thesis, an active control group is included along with a passive control group.

There are also statistical concerns related to WM training studies. First, it has been argued that most training studies are severely underpowered due to small sample sizes (e.g., meta-analysis of Lampit, Hallock, & Valenzuela, 2014; median group size of 22), which increases the risk for both false negatives and false positives (Button et al., 2013). Second, most studies are based on the null hypothesis significance testing, which is not a suitable tool for evaluating the evidence for the null hypothesis or for the lack of evidence for either of the two hypotheses. Null hypothesis significance testing is suitable for evaluating the probability (which is hoped to be low), for the current, or more extreme data in the long run, when there is nothing going on (i.e., assuming there are no training effects).

An alternative to the null hypothesis significance testing is the calculation of the Bayes Factor (BF). BF is the ratio between the likelihood of the data under one hypothesis (e.g., H^1: training has an effect on specific outcome) relative to another hypothesis (e.g., H^0: training has no effect on specific outcome). BFs allow for drawing conclusions about the evidence supporting the presence or the absence of an effect, or whether there is not enough evidence to support either of the two hypotheses sufficiently (Dienes, 2014).

Not many studies have applied BFs to evaluate the effectiveness of WM training, especially in children. Studies applying BF in younger adults (DeSimoni & von Bastian, 2018; Dougherty, Hamovitz, & Tidwell, 2016) and older adults (Guye & von Bastian, 2017), have supported the conclusion of presence of near transfer effects and absence of far transfer effects. For example, by applying BF, Dougherty and his colleagues (2016) demonstrated that studies with passive control groups strongly favored the presence of the effect, but those with active controls moderately favored the absence of the effect. For the present summary, the data of Studies III and IV were reanalysed with Bayesian analysis in order to evaluate the evidence in favor of both the presence and the absence of the effect.

1.5 AIMS OF THE PRESENT THESIS

In sum, in many cognitive tasks, including mathematical school tasks, reasoning and reading comprehension, WM has shown to play a major role. However, it is unsettled why WM is such an important factor underlying individual differences in complex cognitive skills and whether the relationship is causal or not. In the present thesis, I explore the mechanisms explaining the relationship between WM and scholastic skills in the context of WM assessment (Studies I and II), and explore its possible causal underpinning in the context of WM training (Studies III and IV) in children and adolescents. The more specific questions of present thesis are:

1. Given the high correlations between WM span tasks and scholastic skills replicated robustly in previous studies, does natural environmental noise (external distraction) disrupt or facilitate the measured WM span score's ability to predict scholastic skills? (Study I)
2. Does a demand of the processing component (internal distraction) of a WM task contribute the measured WM span's ability to predict scholastic skills and, on the other hand, are there differences related to tasks which are to be predicted? (Study II)
3. Since all components of WM correlate with numeracy, can math scores be improved by computerised training of any component of WM? (Study III)
4. Given the important role of both WM capacity and earlier mathematical skills in later mathematical skills, are group-based training activities

addressed to improve WM and counting or purely counting beneficial in improving children's numerical skills? (Study IV)

In order to answer the research questions, in Studies I and II, complex WM span tasks were constructed and applied together with a broad evaluation of fluid intelligence and scholastic skills (mathematical skills, reading comprehension, reading skills, GPA). In Study I, WM was assessed in a school classroom, in which environmental noise (i.e., external distraction) is inherently present. The noise, for example music from the neighboring classroom, was documented by the research assistants during the WM assessment and classified as speech-noise, non-speech noise or both speech and non-speech noises. Thus, it is possible to evaluate the contribution of noise on WM span and most importantly on WM-scholastic skills relationship.

In Study II, the amount of internal distraction was operationalised as the number of cognitive steps to be performed in a given time in the processing component of the complex span task. The processing component of the task with relatively lower internal distraction, that is, the Reading Span task, required judging a sentence. The processing component with relatively higher internal distraction, that is, the Word Problem Span task, required solving an arithmetic word problem. The amount of words, and the pace of presentation was controlled across the tasks. Since the Reading Span task lacked the arithmetic problem solving, but was in other means similar to Word Problem span task, it was considered having fewer cognitive steps, and thus to produce relatively less internal distraction than the Word Problem span task.

In Study III, the participating children were randomly assigned to six different experimental groups of adaptive WM training: (1) visuospatial STM, (2) visuospatial WM, (3) verbal STM, (4) verbal WM, (5) active controls, and (6) passive controls. In Study IV, the participating children were assigned to three different experimental groups (1) counting training (i.e. domain-specific numerical training); and (2) simultaneous training of WM and counting (i.e. domain-general WM training combined with domain-specific numerical training) and (3) passive control group. The interventions included training of pre-selected tasks together with research assistant two times a week for five (Study III) or four (Study IV) weeks. In both studies, broad battery of WM and numeracy assessments were conducted as pre- and post-training assessments.

It is worth noting, that the age of the participating children, and some of the assessment methods varies across the four studies presented and summarised in current thesis, due to the specific research questions in the original research papers. For example, in Studies I and II participants were school aged children of ten to sixteen years, while in Studies III and IV participants were five to six years old kindergarteners. The sample choice consisting of school aged participants was justified in first two studies, because WM measurement was known to be suspect for attentional manipulation in children of this age group (Lépine et al., 2005). Furthermore, the associations between WM and achievement is well known in school age and thus provide a good opportunity to validate the present results with earlier results. The interventions in Studies III and IV in turn, were conducted for younger children. The practical aim of these studies was to obtain intervention methods that could be applied in real life environments of children's daily routine, that is, in the kindergarten.

Furthermore, the original goals of the original research papers are somewhat different than presented in this summary: in order to summarise the theoretical contribution of the four studies, the theoretically relevant results are highlighted in this summary and some constructs are labelled slightly differently, especially regarding Studies I and II. For example, in the original paper of Study I the main goal of was to evaluate the feasibility of WM tablet tasks in group settings. However, it additionally explored the contribution of environmental noise on performance, which is focused here, since it is relevant in the context of present theoretical discussion. Consistently, new constructs of external distraction (environmental noise, Study I) and internal distraction (processing demand, Study II) are presented in current presentation, in order to contrast the different aspects of attentional demands that were evaluated in Studies I and II.

2 METHODS

2.1 PARTICIPANTS (STUDIES I – IV)

A total of 1069 subjects (493 boys, 576 girls), aged from 5 to 16 years participated the studies. The sample of Study I consisted of 837 early adolescents (376 boys, 461 girls). The mean age of the participants was 12 years (M = 12.3; SD = 0.4) and the ages of the participants ranged from 10 to 14 years. In Study II, the participants were 72 adolescents (28 boys, 44 girls). The mean age of the participants was 16 years (M = 15.9; SD = 0.3). In Study III, the participants were 99 (57 boys, 42 girls) six years old (all children were attending their last two months of kindergarten). In Study IV, the participants were 61 children (32 boys, 29 girls). The mean age of the participants was 6 years (M = 5.9; SD = 0.8).

All participants were native speakers of Finnish and school aged participants were students in the regular classes. The sample of the present thesis was recruited in both small and bigger cities in Finland. In all substudies parents' written consent and, if relevant, student's assent was required for children's participation. The parents were advised to discuss the study with their offspring to ensure their child's own willingness to participate. Teachers of the participating classrooms gave their written consents for the data collections to be conducted during the lessons. All the research plans of the current studies have been evaluated by the local Ethics Committees and the committees have given an ethical statement verifying that the studies do not pose any ethical concerns.

2.2 MATERIALS (STUDIES I – IV)

See Appendix for the availability of the materials, data and the analysis scripts of the present thesis.

2.2.1 COMPLEX SPAN TASKS (STUDIES I & II)

In Studies I and II, WM was assessed with complex span tasks. In Study I the Counting Span task and Reading Span task were used. In Study II the Reading Span task and Word Problem Span task were used. In all complex span tasks, the task started with a block of three trials consisting of two processing and storage tasks (i.e., two items to remember). The amount of processing and storage tasks increased by one in the next block if a participant recalled at least one of the three trials correctly on a given list length. This continued until the participant would not recall any of the three trials correctly, at which point the task would be discontinued. In all complex span tasks, first, there were practice trials. The tasks differed in their presentation mode, which was a mobile tablet device in Study I and a computer in Study II. Furthermore, in Study II, the presentation rate of the tasks was controlled by presenting the processing component word by word.

2.2.1.1 Counting Span task (Study I)

The Counting Span task was modified from the version originally introduced by Case and colleagues (1982). The task consists of counting yellow dots (processing component) presented on a black tablet screen and storing the number of dots in each set in one's memory (storing component). After the presentation of all sets of dots, the participants entered the recalled number of dots on the tablet screen by touching the numbers on the screen in the correct serial order.

2.2.1.2 Reading Span task (Studies I & II)

The Reading Span tasks were modified from the original version by Daneman and Carpenter (1980). The task consisted of reading sentences, judging their trueness (processing component) and storing unrelated written letters (Study I) or words (Study II) in one's memory (storing component).

In Study I, a sentence first appeared on the tablet screen, followed by two boxes with "true" and "false" texts (processing component). After the participant responded by touching the box, an unrelated letter (storing component) appeared. The sentences were short true or false statements, such as "*A rabbit is green.*" Half of the statements were true and the other half false.

In Study II, the to-be-remembered items were concrete Finnish nouns with four to six letters. These memory items were presented before each processing component. They were unrelated to the information presented in the processing component of the tasks. The processing component of the Reading Span task consisted of sentences like "*A boy had red candies and got blue candies from his mother*", which all consisted of ten words in Finnish language (*"Pojalla oli punaisia karamelleja ja hän sai äidiltään sinisiä karamelleja")*. Each sentence was followed by a word to be judged as correct if it was semantically related to the content of the sentences (superordinate semantic category, in this example *"sweets"*) or incorrect if it was semantically unrelated to the content of the sentence (in this example *"animal"*). In order to control for the presentation rate with the Word Problem Span (see below), the sentences appeared on the screen word by word, one word appearing for 1000 ms after a delay of 250 ms.

2.2.1.3 Word Problem Span task (Study II)

The Word Problem Span task consisted of reading word problems, judging their trueness (processing component) and storing unrelated words in one's memory (storing component). The to-be-remembered items were concrete Finnish nouns with four to six letters. These memory items were presented before each processing component. They were unrelated to the information presented in the processing component of the tasks. The processing task of the Word Problem Span task consisted of sentences including a mathematical word problem, half of which required addition and half subtraction, for example *"A girl had 9 apples and she got 5 apples more."*. All sentences had ten words in Finnish language (*"Tytöllä oli 9 omenaa ja hän sai 5 omenaa lisää"*). The operations were taken randomly from a pool of all possible additions and subtractions in which the result fell within a range of 1–20. The number zero was not used in the operations. Each word problem sentence was followed by a correct or an incorrect result and a question mark. Similar to the Reading Span task, the sentences appeared on the screen word by word, one word appearing for 1000 ms after a delay of 250 ms.

2.2.2 COMPREHENSIVE WM BATTERIES (STUDIES III & IV)

The WM tasks differed in their presentation mode, which was a computer in Study III and a paper-and-pencil in Study IV.

In Study III, WM was assessed with the computerised Automated Working Memory Assessment (AWMA; Alloway, 2007), and in Study IV, the Working Memory Test Battery for Children (WMTB–C; Pickering & Gathercole, 2001) was used in assessing WM broadly. Because AWMA or WMTB-C are not standardized in Finland or translated to Finnish, only the visuospatial tasks were administered using original items in the test batteries. The verbal tasks were adapted to the Finnish language on the basis of the original English version of AWMA. In WMTB-C the verbal tasks were replaced with corresponding tasks from WISC-IV.

The AWMA and WMTB-C are based on Baddeley's (1986, 2000) WM model, assessing the different WM components with several tasks. In both of the test batteries, the tasks are assumed to reflect STM, if no additional processing is required in the task and WM, if there is a processing component along with memory component. Furthermore, the tasks are either verbal or visuospatial in nature. Thus, in Studies III and IV, tasks were assumed to assess visuospatial STM and WM and verbal STM and WM. In all tasks the list length increases, until the participant fails to respond in the memory component, and the span is thus assumed to reflect the amount of information the participant can hold in memory.

2.2.2.1 Visuospatial STM

In the **Dot matrix** task (Studies III and Study IV) the participants were presented with a sequence of red dots on a 4 × 4 grid for two seconds each. The presentation was on the computer screen in Study III and on paper cards in Study IV. The child was required to point with a finger to the positions of the dots that had appeared in the order of the appearance. In the **Block recall** task (Study III), a sequence of cubes was highlighted on a screen with nine randomly located cubes. The child's task was to repeat the sequence in the same order by pointing at them with a finger.

2.2.2.2 Visuospatial WM

In the **Odd one out** task (Study III and Study IV), the participants were presented with a row of three shapes and instructed to point with a finger to the odd one out and remember its location. At the end of the task, the child was instructed to recall the position of the shape that he or she had identified as being different. In the **Mister X** task (Study III), the child was presented with a picture of two people, each of whom was holding a ball in one of their hands. One of the persons was rotated and the child's task was to tell if the two people were holding the ball in the same hand and to point the location of the rotated person's ball.

2.2.2.3 Verbal STM

In the **Word span forward** task (Study IV), the participants were instructed to recall orally series of word lists in the correct serial order. In the **Non-word** task (Study IV), the procedure was the same as for the word span forward, except that it consisted of lists of non-words. The **Digit span forward** task (Study III) was presented and scored as recommended in the WISC-III Manual (Wechsler, 2010). The participant was required to recall a list of auditorily presented digits (1–9) in a correct order. The lists were presented in ascending order, two lists in each set, starting from lists of two digits and continuing to lists of nine digits. The task was continued until the participant made a mistake in both of the two lists of a set.

2.2.2.4 Verbal WM

In the **Listening span** task (Study III), the participants were presented with a set of spoken sentences and instructed to judge whether a sentence was true or false by answering orally and retaining the final word of the sentence in sequence. In the **Counting span** task (Study III), the children were presented with a visual array of red circles and blue triangles. They were instructed to count the number of circles by pointing at them with a finger, and to recall the tallies of the circles after the array disappeared. In both tasks, the children responded orally.

In the **Word span backwards** task (Study IV) the procedure was similar to the Word span task with the exception that in this task the participant had to recall the presented words in reverse order. The test began with a sequence of two words, and it ended when the participant recalled two out of the four trials for a certain sequence incorrectly. The **Digit span backwards** (Study IV) was conducted similarly to the digit span forward, except that the participant was required to recall the digits in reverse order from what was presented.

2.2.3 SCHOLASTIC SKILLS (STUDIES I–IV)

2.2.3.1 Mathematical skills

Basic arithmetic test (Aunola & Räsänen, 2007; Räsänen, Salminen, Wilson, Aunio, & Dehaene, 2009) was used to assess fluency in arithmetic skills in Study I. Basic arithmetic test is a speeded group-administered test, in which each participant is required to complete as many arithmetic operations as possible within a three-minute time limit. The sixth-grade form of the test consists of 10 additions and 11 subtractions. Seven tasks include both additions and subtractions or multiplication and division problems to be solved.

The **National curriculum test in mathematics** (Mathematics NC) was used for assessing mathematical skills in Study II. The test, developed and validated by Finland's National Board of Education, included three subtests: Mental Arithmetic (α = .77), Multiple Choice (α = .85) and Problem Solving tasks (α = .88) (Hirvonen, 2012). The mean of the national sample of 4929 students in 2011 was 43.5 (SD = 17.15; Hirvonen, 2012). The sample mean of the Mathematics NC, M = 46.66 (SD = 13.97) in Study II did not significantly differ from the mean of the national sample [$t(67)$ = 0.01, p = ns.].

The **Diagnostic Tests for Metacognition and Mathematics test battery** (Salonen et al., 1994) was used for assessing counting skills in Study III. Eight verbal counting subtasks were adopted from the test. The tasks require counting up to 50, counting the number of objects, counting up from a number, counting up a specific amount from a number, counting backward from a number, and counting backward from one number to another.

The **Early numeracy test** (Van Luit, Van de Rijt, & Aunio, 2006) was adopted for assessing mathematical skills in Study IV. Three subscales of version A were used: (1) use of number words (counting forwards and backwards up to 20, using cardinal and ordinal numbers); (2) structured counting (counting while pointing to objects, recognizing numbers on a die); and (3) resultative counting (counting without pointing to objects).

2.2.3.2 Reading Comprehension

A modified Finnish version of **the Salzburg Reading Test** (Landerl, Wimmer, & Moser, 1997; translated into Finnish by Sini Huemer) was adopted in Study I. It consists of 36 sentences that are each either true or false. The participants were instructed to read each sentence and mark it as either true or false. The time limit for each task was 90 seconds. The score was the number of correctly judged sentences.

A standardized **reading comprehension test** (Nevala, Kairaluoma, Ahonen, Aro, & Holopainen, 2006) was used in Study II. It is commonly used in Finland for clinical screening purposes. Participants read a text, and answered multiple choice questions. The text was two pages long and concerned the use of technology.

2.2.3.3 Reading Fluency

Word chain task (Holopainen, Kairaluoma, Nevala, Ahonen, & Aro, 2004) was adopted in the Studies I and II. In this task there were 25 word chains, four words in each chain written without spaces between them (e.g., *nicecrashcoatsnail*). The participants task was to separate the words correctly in 90 seconds by a vertical line drawn with a pencil. The words in the chains were semantically unrelated.

2.2.3.4 Spelling

Spelling task in which the participant had to find spelling errors among one hundred written words (Holopainen et al., 2004) was adopted in Study I. The participants' task was to mark (with a vertical line) as many spelling errors as possible within 3.5 minutes. The possible types of spelling errors were: a missing letter, a wrong letter, or an extra letter. The score was the number of correctly identified errors.

2.2.3.5 GPA

Participants' self-reported grade point average of the previous spring semester was collected in Study I. The participants were asked to report their grade point average from the previous spring semester. Children's self-reported school grades have been shown to correlate .86 with their actual grades from the school registers (Ahonen & Kiuru, 2013-2007).

School reports of the current term were obtained to collect the grade point average in Study II. Overall school achievement was determined with the Finnish Grade Point Average, which consists of the average of grades for Finnish, English, Swedish, Mathematics, Biology, Geography, Physics, Chemistry, Religion/Life Philosophy, History, Social Studies and Health Education.

2.2.4 INTELLIGENCE (STUDIES I–IV)

Raven's test was used for assessing fluid intelligence in the Studies I–III. In Studies I and II, Raven Standard Progressive Matrices (Raven, Raven, & Court, 1998) and in Study III Raven's Coloured Progressive Matrices (Raven, Court, & Raven, 1995) were adopted. Raven's test consists of diagrams with one part missing. The participants are asked to select the correct part that would complete each design, and the test increases in difficulty. In Study I, only half of the items were used and alternating items were selected to be presented. In Study II and Study III, all items were presented. Two subtests of **WISC-III** (Wechsler, 2010) (Block design and vocabulary task) were adopted in Study IV for assessing crystallized intelligence. These subtests were selected because they have high reliability and high correlation with the fullscale IQ (Silverstein, 1982).

2.2.5 ENVIRONMENTAL NOISE IN THE CLASSROOM DURING WM ASSESSMENT (STUDY I)

During the WM testing sessions, two research assistants were instructed to document and describe all environmental distractions, noise and other sudden sounds that they observed. The observing and documenting of the environmental noise were done over the total session in which both tasks were administered. Afterwards, the verbal descriptions were coded by another researcher according to four categories: "no noise," "speech noise," "non-speech noise," and when applicable, "both speech and non-speech noise." The noise was scored as speech noise when verbal auditory noise was reported, for example the teacher (in nine sessions) or students (in 11 sessions) asked something or the central radio announcement started (in two sessions). The noise was scored as non-speech noise when non-verbal auditory noise was reported, for example the school bell rang (in five sessions) or a piano was being played in the neighboring classroom (in two sessions). However, in some cases the nature of the noise was not that clear, such as unspecified sounds from the school's corridor (in five sessions) or from the neighboring classroom (in two sessions). In above-mentioned cases, the noise was scored as non-speech noise.

2.2.6 THE INTERVENTIONS (STUDIES III & IV)

2.2.6.1 Computerised WM intervention (Study III)

Four computerized WM training programs were developed for use of Study III. Common to the developed training programs was the adaptiveness of the task difficulty, which was matched to the actual performance of each child. In the **visuospatial STM training**, the child was presented with a matrix in which animal figures appeared for two seconds in half of the squares. The child was instructed to recall in which squares of the matrix the animals had appeared and to use a mouse to click on the squares on an empty matrix after their disappearance. In the **visuospatial WM training**, the child was presented with three animals located above places to hide (e.g., a tree or a stone). One of the animals was different from the others. The child was instructed to point out which one was different. After that, the animal hid (moved behind the element). The child was expected to retain the position of the hiding places and a new set of animals and hiding places appeared. At the end of the sequence the child was instructed to recall the hiding places in their presentation order by mouse click. In the **verbal STM training**, the child was instructed to learn new words that animals taught. One animal appeared on the screen and the non-word was heard on the headphones. The child was instructed to repeat orally the non-words in the correct order. In the **verbal WM training**, the child was instructed to learn new words from fish and to select the fish that had taught the word on the basis of a given description. In this training two fish appeared on the screen and a syllable emerged from the headphones (storage component). After that the child was instructed to point with a finger at the fish that had taught the word (processing component). At the end of a series, the child orally recalled the syllable/syllables in the order of their presentation and the next trial began.

2.2.6.2 Group interventions (Study IV)

In Study IV the training was conducted in groups of 4 to 7 children. In the **Counting training** group, children practiced a variety of counting-based activities at a group. For example, children practiced counting (number word sequence; forward, backward) starting with 1 to 10 and increasing the sequence in each week of training. They played a bingo game in which they practiced connecting number symbols with the same number of dots, and practiced number word sequence by walking on a digit path. They also played a game in which they had to compare number symbols and a set of dots. In the last week of the training, the activities contained digits from 1 to 100.

In the **WM and counting training** group, children practiced a variety of memory and counting based activities at a group. For example, in the first week of training, the numbers from 1 to 10 were practiced in a game, in which the children had to remember things that they could bring on a holiday, and how many they could bring, for example: ‘I go on a holiday and I take one toothbrush with me.’ The number of different items and their counts increased after every child’s turn. The children also played a memory game in which they had to find right pairs by finding a certain card with a number symbol on it and a card with corresponding amounts of dots.

2.3 PROCEDURES (STUDIES I-V)

In both Studies I and II, the assessments were conducted in schools during typical school hours. In Study I, the assessments were conducted at groups in classrooms. A small subsample was assessed individually. In Study II all WM assessments were conducted individually. In both studies, the scholastic skills and fluid intelligence were assessed in the school classes.

In Studies III and IV, the pre- and post-assessments and the training sessions were conducted in kindergarten during the regular kindergarten activities. The pre- and post-assessments and the training of Study III were conducted individually. The training of Study IV was conducted in groups of children.

3 RESULTS AND DISCUSSION

The WM span tasks in Studies I and II were analyzed using partial credit unit (PCU) scores (Conway et al., 2005). In this scoring method, the mean number of correctly recalled memory items within a list length is calculated, and these proportions are then averaged to obtain the score. This scoring method is recommended by Conway and his colleagues (2005) based on solid internal consistency. For all the analyses using null hypothesis significance testing, the criterion for statistical significance was set a priori at $\alpha = .05$.

The data of Studies III and IV were first analysed with traditional null hypothesis significance testing (NHST) in the original published articles. However, to evaluate the statistical support for both null hypothesis (H^0: no training effects) and alternative hypothesis (H^1: training effects) the post-hoc Bayesian analysis were later conducted for both data of Study III and IV, and are reported in the present summary together with NHST. The Jeffrey-Zellner-Siow (JZS) Bayes factors were computed for each cognitive test with default prior scales (Rouder, Morey, Speckman, & Province, 2012) using the "bayesfactor" package (Morey, Rouder, Jamil, & Morey, 2015). BFs below 1 represent evidence for the H^0 (no training effect), and BFs above 1 indicate evidence in favor of the H^1 (training effect). BF of 1 reflects perfect ambiguity (i.e., the data support both hypotheses equally) (see Table 1). For example, a BF of 10 in favor of the H^1 means that the data are ten times more likely under H^1 than H^0.

Table 1. *The interpretation of Bayes Factors (Jeffreys, 1961; modified by Lee & Wagenmakers, 2013)*

Bayes factor BF_{10}	Interpretation
> 100	Extreme evidence for H^1
30-100	Very Strong evidence for H^1
10-30	Strong evidence for H^1
3-10	Moderate evidence for H^1
1-3	Anecdotal evidence for H^1
1	No evidence
1/3-1	Anecdotal evidence for H^0
1/10-1/3	Moderate evidence for H^0
1/30-1/10	Strong evidence for H^0
1/100-1/30	Very Strong evidence for H^0
<1/100	Extreme evidence for H^0

3.1 STUDY I

To investigate the associations of WM and scholastic skills, first the correlations between the two WM tasks (Counting Span task and Reading Span task), fluid intelligence and scholastic skills (arithmetical skills, reading comprehension, reading fluency, spelling and GPA) were studied. Both WM span tasks correlated positively with scholastic skills and fluid intelligence (Table 2). The correlations were generally similar in strength compared to those obtained in vast amount of earlier studies evaluating the associations between complex span tasks and scholastic skills (Alloway & Alloway, 2010; Gathercole et al., 2004; Lépine et al., 2005; Seigneuric & Ehrlich 2005).

Table 2. *Summary of the correlations between WM span tasks and scholastic skills in Study I (N=873)*

	1	2	3	4	5	6	7
1. Counting Span	--						
2. Reading Span	.44						
3. Fluid Intelligence	.42	.36					
4. Arithmetic Skills	.45	.34	.40				
5. Reading Fluency	.43	.40	.27	.47			
6. Spelling	.42	.44	.26	.43	.75		
7. Reading Comprehension	.27	.36	.18	.39	.65	.69	
8. GPA	.42	.36	.38	.35	.43	.44	.40

Note. All correlations are statistically significant, $p < .001$.

Next, to investigate the contribution of natural environmental noise in WM, the memory scores of WM span tasks and the correlations between WM and scholastic skills under natural environmental noise was studied. The types of environmental noises that the examiners reported in the classrooms during the WM span assessment sessions were classified as follows: no noise (28 sessions, 397 participants, mean of the number of participants in a session = 19.54; SD = 4.51; median = 20), speech noise (12 sessions, 174 participants, mean of the number of participants in a session = 22.89; SD = 4.30; median = 24), nonspeech noise (14 sessions, 180 participants, mean of the number of participants in a session = 18.68; SD = 3.81; median = 19), and simultaneous speech and non-speech noises (9 sessions, 86 participants, mean of the number of participants in a session = 22.23; SD = 4.76; median = 22.00).

3.1.1 PERFORMANCE IN THE WM SPAN TASKS UNDER ENVIRONMENTAL NOISE

The mean scores, standard deviations and the reliability estimates for the WM performances in the span tasks, separately for each type of noise group, are reported in Table 3. The analysis shows that the performance levels (PCU scores) did not differ by type of noise group, that is, those exposed speech-based or non-speech-based noise or no exposion to noise during WM assessment (in Counting Span task: $F(833, 3) = 0.92$, $p = .43$, $\eta^2 = .003$; in Reading Span task: $F(833, 3) = 0.57$, $p = .64$, $\eta^2 = .002$).

In order to obtain the reliability estimates for WM span tasks, the first presentation of all the sets of different lengths into a single PCU score, the second presentation into a single PCU score, and the third presentation into a single PCU score were combined. These three subscores for each WM task were used to compute Cronbach's α as a measure of reliability (Conway et al., 2005; Engle et al., 1999). The reliability estimates are reported in Table xx, and they demonstrated high internal consistency for both WM span tasks in all noise groups. This indicates that participants who responded with the correct answer for one set of span tasks tended to respond with the correct answer also on the other sets.

Table 3. *The mean scores, standard deviations and the reliability estimates for the WM performances in the span tasks*

		No-noise (*N* = 397)	Speech noise (*N* = 174)	Non-speech noise (*N* = 180)	Both types of noises (*N* = 86)
Counting Span (PCU)	Mean (SD)	0.45 (0.19)	0.47 (0.19)	0.44 (0.19)	0.46 (0.20)
	Cronbach's *α*	0.95	0.95	0.95	0.96
Reading Span (PCU)	Mean (SD)	0.61 (0.18)	0.63 (0.18)	0.61 (0.18)	0.60 (0.22)
	Cronbach's *α*	0.94	0.91	0.92	0.95

3.1.2 ASSOCIATIONS BETWEEN THE MEASURED WM SPAN AND COGNITIVE SKILLS UNDER EXTERNAL DISTRACTION

Next, the associations between the measured WM span and scholastic skills were evaluated under external natural speech- and non-speech noise. The hierarchical linear analyses suggested that although speech noise or non-speech noise alone did not moderate the relationship between the WM span scores and the achievement scores, the groups with both noises had in some cases higher correlations compared to the other groups (Figures 1 and 2). More specifically, in the group with both noises the interaction effect was present in the relationship between the Counting Span and Raven's test, Counting Span and Arithmetic skills, and Counting Span and two Reading skills tests, as well as between the Reading Span and Raven's test (Figures 1 a-d and 2a).

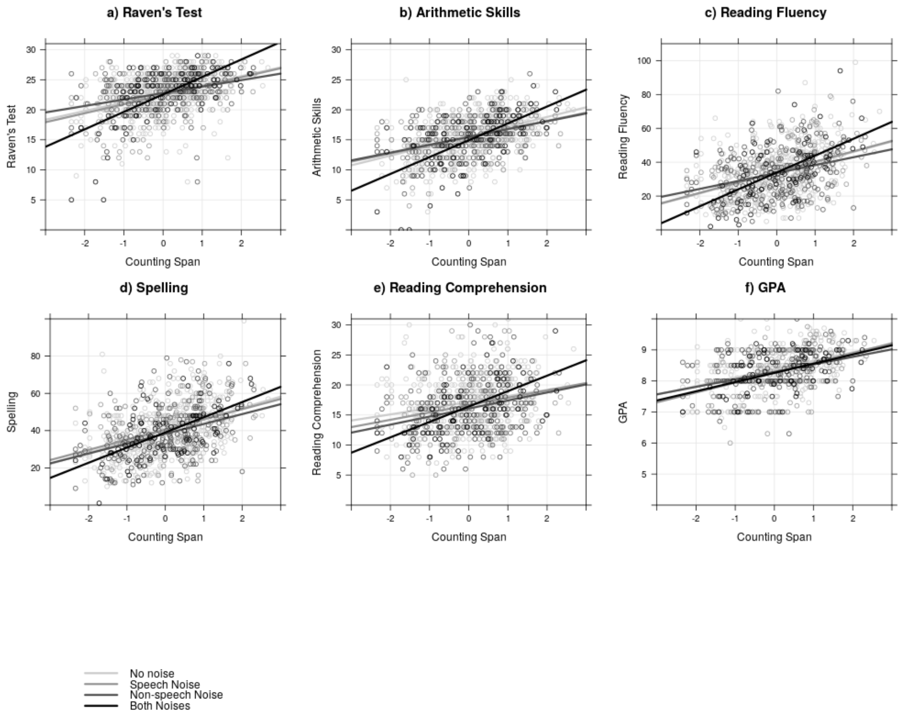

Figure 1 The correlations between Counting Span and scholastic skills in noise groups (Study I)

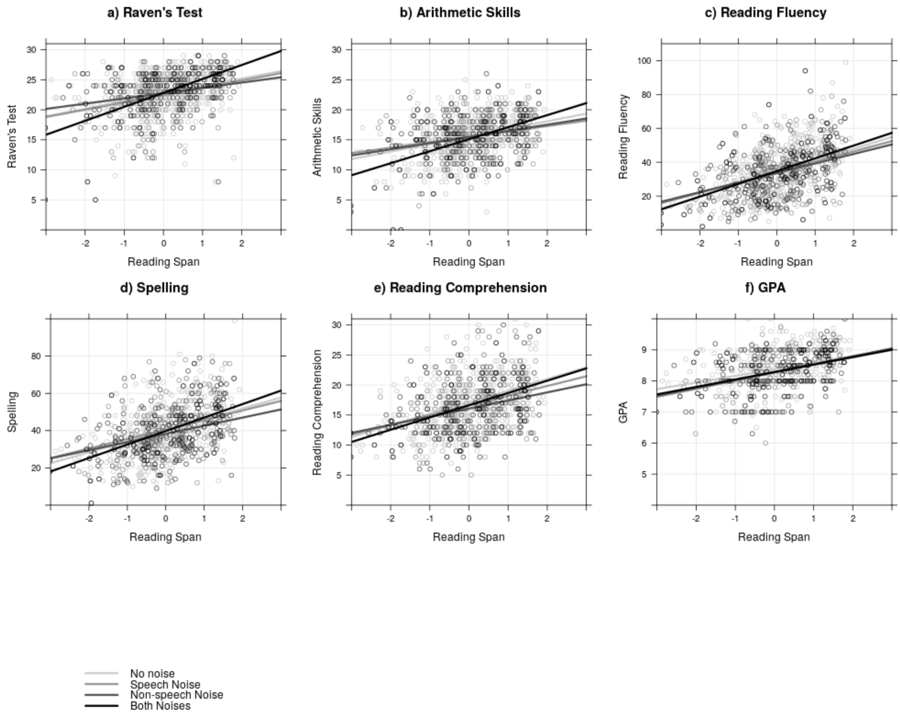

Figure 2 The correlations between Reading Span and scholastic skills in noise groups (Study I)

Taken together, the results of Study I suggest that external distraction have a role in the relationship between WM and cognitive abilities. First, the external distraction, in this case the environmental noise, contributed to the predictive utility of the WM tasks making the WM task in certain cases more predictive, not less predictive of the cognitive skill. More specifically, the correlations were higher in classrooms with both speech and non-speech noise compared to classrooms with only speech or non-speech noise or no noise, when Counting span was adopted in predicting Raven's test, arithmetical skills, spelling and reading comprehension. Similar result emerged when Reading span was adopted in predicting Raven's test.

These results are in line with and extend those of a study by Sullivan, Osman, and Schafer (2015), who found there to be a stronger relation between WM span and comprehension in school-aged children when these were exposed to noise compared to when not. Because the occurrences of noise was not controlled for, it is worth noting that these results might reflect the amount of noise rather than the type of noise, or there can be some other uncontrolled aspects that were related to those particular school classes. However, since the WM span tasks' reliability estimates were similar regardless of the environmental noise in both WM tasks, the tasks' ability to consistently catch the individuals' performance levels could be considered reliable even under environmental noise.

3.2 STUDY II

Next, to investigate the contribution of internal distraction in WM, the memory scores of two WM span tasks varying in their processing demand were compared with each other and their associations with intelligence and scholastic skills was studied. Again, both WM span tasks correlated positively with scholastic skills and fluid intelligence (Table 4).

Table 4. *Summary of the correlations between WM span tasks and scholastic skills in Study II (N=68)*

	1	2	3	4	5	6	7
1. Reading Span							
2. Word Problem Span	.62***						
3. Raven's test	.24*	.34**					
4. Mathematics NC	.38**	.39***	.50***				
5. Reading Comprehension	.18	.22	.29*	.27*			
6. Word Chains	.15	.40	.17	.21	.32**		
7. Mathematics Grade	.35**	.33**	.46***	.81***	.16	.30*	
8. GPA-F	.38***	.30*	.37**	.71***	.41***	.45***	.87***

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

3.2.1 PERFORMANCE IN THE WM SPAN TASKS UNDER TASK'S INTERNAL DISTRACTION

First, the results indicated that the mean span scores were significantly higher in the Reading Span task (lower demand) than Word Problem Span task (higher demand) $t(67) = 2.72$, $p < .01$; *95% CI* 0.126, 0.019; $d = .29$. When comparing the performance in the processing component, the response times did not differ between the two tasks $t(67) = .73$, ns., but the accuracy was significantly higher in the Reading Span task than Word Problem Span task $t(67) = 9.29$, $p < .001$; *95% CI* 0.074, 0.048; $d = 1.41$.

3.2.2 ASSOCIATIONS BETWEEN THE MEASURED WM SPAN AND COGNITIVE SKILLS UNDER TASK'S INTERNAL DISTRACTION

Next, to determine if the more demanding WM span task accounted for any additional variance in the higher order cognitive functions (fluid intelligence, academic skills, and school achievement) beyond the less demanding WM span task, two hierarchical regression analyses were conducted. Two WM span tasks were entered as independent variables, one at a time in both orders (Model A and Model B), and the proportion of explained variance in each higher order cognition measure separately (univariate regression analyses) was studied (see Table 5). When the Word Problem Span task was entered into Model A as the first variable, it predicted a significant amount of variance (9–15 %) of all higher-level skills: Raven's test, Mathematics NC and GPA. Entering the Reading Span task next did not significantly increase the amount of explained variance, except for the GPA (6 %). In Model B the Reading Span task was entered into the regression model first. In this case the Reading Span task explained a significant amount of variance (6–14 %) in the higher-level tasks. However most critically, entering Word Problem Span task to the model as the second independent variable significantly increased the amount of explained variance in Raven's test (6 %), but not in the scholastic performance variables.

Table 5. *Hierarchical regression analysis for predicting fluid intelligence and scholastic performance with two working memory span tasks (N = 68)*

	Raven's test		Mathematics NC		GPA-F	
Model A	*ΔR^2*	*ΔF*	*ΔR^2*	*ΔF*	*ΔR^2*	*ΔF*
Stage 1. Word Problem Span	0.12**	8.7	0.15***	11.76	0.09*	6.67
Stage 2. Reading Span	0	0.13	0.03	2.43	0.06*	4.61
Model B						
Stage 1. Reading Span	0.06*	4.19	0.14**	10.98	0.14***	11.13
Stage 2. Word Problem Span	0.06*	4.32	0.04	3.12	0.01	0.58

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

3.2.3 EXTRACTING SHARED AND UNIQUE VARIANCE WHEN PREDICTING SCHOLASTIC SKILLS

Because the correlations between the measured WM span and other cognitive tasks were rather high, as is typical (positive manifold), the correlations and regression analyses do not reveal the unique variance explained by the different kinds of WM tasks. For that reason, in Study II the regression commonality analysis was used in order to handle the multicollinearity and to decompose the unique and common variance between the measures (Nimon & Reio, 2011). In other words, the proportion of explained variance that is unique to each WM task was calculated. The three separate commonality analyses, in which dependent variables were Raven's test, Mathematic NC and GPA revealed that 49 %, 62 % and 55 % of the total explained variance was common to Word Problem Span and Reading Span, respectively (see Figure 3). However, in Raven's test, 50% of the total explained variance was unique to Word Problem Span and 2% to Reading Span, whereas in GPA, 5% of total explained variance was unique to Word Problem Span and 40% to Reading Span. In Mathematics, NC 22% of the total explained variance was unique to the Word Problem span and 17% to Reading Span.

In sum, these results describe first of all that a substantial amount, about 50–60 %, from the variance that the complex span tasks explain the scholastic skills is shared between two complex span tasks. However, also the performance in particular WM tasks explains a substantial amount of unique variance in particular criterion, that is, the Word Problem Span explained performance in Raven’s test and Reading Span explained GPA. Thus, the amount of unique contribution is not specific to the particular WM span task, but it differs across the outcomes.

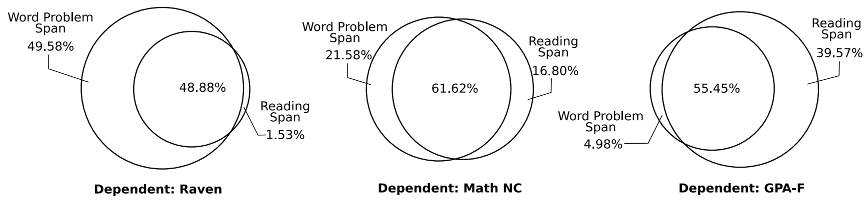

Figure 3 Venn diagrams representing the common and unique variance (%) of the two WM span tasks (Word Problem Span and Reading Span) when predicting three different higher order cognitive measures (Dependent; Raven: Raven’s Progressive Matrices; Mathematics NC: National Curriculum Test for Mathematics; GPA-F: Grade Point Average, Finland)

Together, these results indicate that the increasing demands of the WM processing tasks contributes both the memory scores and to the associations between WM and cognitive skills. However, the amount of unique contribution of processing demand is not specific to the particular WM span task, but it differs across the cognitive tasks that are being predicted. First of all, the processing demand do not affect the relationship between performance in the WM span task and scholastic performance, namely performance in the Mathematics NC or GPA-F. In fact, the less demanding task predicted variance in the GPA-F and the more demanding span task failed to add the explained variance (Lépine et al., 2005; Magimairaj & Montgomery, 2012). However, the relatively more demanding WM span task accounted for an additional part of the variance in fluid intelligence that was not accounted for by the relatively less demanding WM span task. In line with Bunting (2006), this result suggests that a WM task with a more demanding processing task is also a good predictor of higher order cognition, but for a different reason than in the case of a less demanding WM task: it seems to share with fluid intelligence different processes than the less demanding task shares.

3.3 STUDY III

In order to address the domain-specific aspects of WM training, each domain of the WM (verbal STM and WM, visuospatial STM and WM) were addressed by training WM components in separate groups and evaluating the effects of training on 1) corresponding WM components, 2) across WM components and 3) transfer to more complex skill, in this case, the numeracy.

3.3.1 THE NEAR AND FAR-TRANSFER EFFECTS OF TRAINING OF SPECIFIC WM SUBCOMPONENTS

In Study III, the effects of the computerised intervention of verbal and visuospatial subcomponents of WM on 5 to 6 years children's WM (near-transfer) and numeracy (far-transfer) were compared with passive and active control groups. The groups were visuospatial short-term memory (N = 15), visuospatial working memory (N = 16), verbal short-term memory (N = 17), verbal working memory (N = 14), active controls (N = 15), and passive controls (N = 17). Children in the passive control group took part in the regular kindergarten activities and pre- and post-training assessments. In the reanalysis of the data of Study III, the dataset with complete cases (N=69) is used. The imputed data, which appeared to show similar results, are reported in the orginal article.

First of all, the one-way ANOVA was used for evaluating whether the groups differ in pre-assessment. Results of Study III indicate no significant differences in performance between groups at the pre-assessment stage on any of the ten pre-assessment measures, (all F-values < 1.01; p-values non-significant; BFs range 0.08-0.20) meaning strong to moderate evidence for H^0). Next, the near transfer of WM training was investigated with two-way repeated measures ANOVAs between pre- and post-assessment (Time) and Group (4 WM training groups, 2 control groups) scores for each of the WM tasks. Regarding the near transfer outcome variables (verbal and visuospatial STM and WM tasks), the results showed no significant Time × Group interaction (all F-values < 2.52, p-values non-significant) (Table 6). BFs comparing interaction model to the main effects model range 0.06-0.37 meaning that there was more support for the main effect model and it was ranging from anecdotal to strong depending on the WM measure. These results mean that the change in the performance of any of the WM tasks, either in the trained WM task or other WM tasks, did not differ across the six groups from pre-assessment to post-assessment. No further analyses were thus conducted.

When far transfer (transfer to numeracy) was evaluated, a similar pattern of results emerged. There was no Time x Group interaction (F = 0.31, BF = 0.08; strong support for the main effect model). Taken together, these results can be interpreted suggesting that for WM or early numeracy skills, the training was not beneficial, was it verbal or visuospatial, and was it addressed to STM or WM.

Table 6. *Statistics from repeated measures ANOVA's testing the Time x Group interaction effect in Study III (N=69)*

Outcome	*F* for Time x Group interaction	*BF* for Time x group interaction vs. main effects	Interval for *BF*	Verbal label of *BF*
Word span forward	0.31	0.06	±21.4%	Strong for H^0
Non-word span	1.33	0.34	±6.05%	Anecdotal for H^0
Listening span	1.37	0.37	±6.32%	Anecdotal for H^0
Dot matrix	0.46	0.10	±7.54%	Strong for H^0
Odd one out	0.68	0.13	±8.13%	Moderate for H^0
Counting span	0.50	0.10	±2.26%	Strong for H^0
Block recall	0.96	0.18	±3.35%	Moderate for H^0
Mister X	1.44	0.34	±12.17%	Anecdotal for H^0
Numeracy	0.31	0.08	±3.67%	Strong for H^0

In sum, when focusing training on separate components of WM, that is, on verbal and visuospatial domains of simple STM and complex WM functions in Study III, there were no evidence of effects of interventions on any aspects of WM or on counting skills compared to the active or passive control groups. The Bayesian evidence supported the results obtained with the traditional NHST. Most importantly, there were no evidence for training effects in WM training groups transferring to counting skills.

3.4 STUDY IV

Second, in Study IV the effects of training of domain-general WM was compared to the training of the domain-specific training, that is, training the task similar to the domain of interest. Thus, in Study IV, to improve the children's early mathematical skills the WM was trained in one group (WM+Counting training; $N = 23$) and counting (Counting training; $N = 21$) in another group. These groups were compared to the control group ($N = 17$).

First of all, the one-way ANOVA show that in Study IV there were no significant differences in performance between groups at the pre-assessment stage on any of the measures (all F-values < 1.01; p-values non-significant; BFs 0.14-0.61) meaning moderate to anecdotal evidence for H^0. Only for odd-one-out task the evidence was anecdotal for H^1; BF=1.2. Next, the transfer to WM span tasks (i.e., word span forward and backward, digit span forward and backward, odd-one-out and matrix task) of Counting training and WM+Counting training was investigated by calculating repeated measures

ANOVAs for each of the WM tasks using Time (pre- and post-assessment) and Group (Counting training, WM+Counting training, Controls) as factors.

Again, regarding the transfer to WM span tasks, the results showed no significant Time × Group interaction (all F-values < 2.52, p-values non-significant) (Table 7). BFs comparing interaction model to the main effects model range 0.14–0.50 meaning that the data was supported by the main effect model and the strength of the support was ranging from anecdotal to moderate depending on the WM measure. This means that the change in the performance of any of the WM tasks did not differ in the three groups from pre-assessment to post-assessment. No further analyses were thus conducted. However, when far transfer (transfer to numeracy) was evaluated, results showed a significant Time × Group interaction ($F = 3.51$, $p < .05$; BF = 1.65). The BF of 1.65 indicates that data occured 1.65 times more likely under H^1 (training has an effect) than under H^0 (training has no effect) given the priors assumed in the model.

Planned pairwise comparisons showed that the performance gain in Early Numeracy Test in the Counting training group was significantly larger than in the controls ($F[1,58] = 4.21$, $p = .04$, $p^{SID} = 0.13$; $\eta^2 = 0.07$; *Note.* p^{SID} means Sidak corrected p-value). Furthermore, the gain was significantly larger in the Counting group than in the WM+Counting group ($F[1,58] = 6.06$, $p = .02$, $p^{SID} = .05$; $\eta^2 = 0.10$). In contrast, the gain in the WM+Counting group did not differ from the controls ($F[1,58] = 0.05$, $p = .82$, $p^{SID} = .99$). This result indicates that the Counting group scores in Early Numeracy Test increased due to training, while the Control group or the WM+Counting group did not show such a pattern as a function of time. This result shows that the Counting group scores in Early Numeracy Test increased due to training, while the Control group or the WM+Counting group did not show such a pattern as a function of time. Although it remains unclear if combined WM and counting training has effect on WM or counting, these results can be interpreted suggesting that at least for early numeracy skills, the training of the task-specific domain of interest is beneficial.

Table 7. *Statistics from repeated measures ANOVA's testing the Time x Group interaction effect in Study IV (N=61)*

Outcome	*F* for Time x Group interaction	η^2 for Time x Group interaction	BF for Time x Group interaction vs. main effects	Interval for BF	Verbal label of BF
Word span forward	0.89	0.03	0.27	±4.23%	Moderate for H^0
Word span backwards	0.26	0.01	0.17	±2.51%	Moderate for H^0
Digit span forward	1.79	0.06	0.5	±4.74%	Anecdotal for H^0
Digit span backwards	0.18	0.01	0.15	±2.39%	Moderate for H^0
Odd one out	1.73	0.06	0.49	±3.2%	Anecdotal for H^0
Dot matrix	0.17	0.01	0.14	±5.36%	Moderate for H^0
Numeracy (total)	3.51*	0.11	1.65	±4.36%	Anecdotal for H^1

Note. *=the interaction is statistically significant, $p < 0.05$.

In sum, in Study IV the training was addressed to the domain of interest, in this case, the numeracy. Furthermore, in order to develop methods for purposes of kindergarten groups, the training was placed in a group situation. Numerical skills were targeted in one training group for studying the effects of domain-specific training. The results showed evidence for positive effects of domain-specific counting training on numerical skills. The Bayesian analysis supported the results obtained with NHST, although revealing that the evidence is anecdotal. However, the combined WM and counting training was not beneficial for the WM skills or counting skills. The Bayesian evidence against the interaction effect was strong or anecdotal as its best, depending on the criterion task.

4 GENERAL DISCUSSION

The main question of the present thesis was: what is the nature of mechanisms that explain the relation between WM and scholastic skills in children and adolescents. In particular, do attentional requirements strengthen the relationship, and does improving of WM enhance performance in scholastic skills, which would indicate that the relationship is causal. The first set of questions in Studies I and II concentrated on investigating the contribution of demands of WM assessment on the relationship between WM and cognitive skills. More specifically, Study I evaluated whether natural environmental noise (external distraction) would disrupt or facilitate the measured WM span score's ability to predict scholastic skills. Study II evaluated whether the demand of the processing component (internal distraction) of a WM task would affect the measured WM span's ability to predict scholastic skills and, on the other hand, whether there are differences related to tasks which are to be predicted. The second set of questions, in Studies III and IV tried to reveal whether the training of domain-general WM capacity and domain-specific counting skills, or both, would affect WM and early mathematical skills.

4.1 WM UNDER DISTRACTION

Together Studies I and II demonstrated, that in order to contribute to the relationship between WM task and cognitive task the distraction can origin either from the environment or from the task itself. First, in Study I, the external distraction, in this case the environmental noise, made the measured WM task performance more predictive, not less predictive of the cognitive skill. Second, in Study II, the internal distraction also improved the correlations between WM task and Raven's test when a WM task with relatively higher processing load was contrasted with WM task with relatively lower processing load. However, the results also demonstrated that the constructs of WM, fluid intelligence and scholastic skills are overlapping and the contribution of distraction show different patterns across cognitive tasks.

These results can be interpreted to support the view that the circumstances which pose high load on controlled attention can help in manifesting the individual differences in WM capacity more clearly (Sörqvist, Stenfelt, & Rönnberg, 2012). Previous findings have shown that individuals with high WM are able to resist interference better than individuals with low WM (Sörqvist et al., 2012). It has been proposed that individuals with high WM are less susceptible to distraction because they can focus (or constrain) their attention better regarding attended targets (Heitz & Engle, 2007), or they have a superior inhibition capacity (Lustig, Hasher, & Zacks, 2007) or because they manage to maintain the goal-directed task set in their working memory even when challenged by stimuli that capture part of their attention (Unsworth & Engle, 2007). It is also possible that individuals with low WM have a more vulnerable rehearsal process, demonstrated by greater irrelevant speech effects when a task's requirements for rehearsal increase (Elliott et al., 2016). However, this relation may not be based only on executive attention (Engle, 2001) but also on other processes related to interference (e.g. Oberauer, Farrell, Jarrold, Pasiecznik, & Greaves, 2012), on activating information in long-term memory (Mogle, Lovett, Stawski, & Sliwinski, 2008; Shelton, Elliott, Matthews, Hill, & Gouvier, 2010; but see Unsworth, 2010) or on the task-specific aspects that a WM span task shares with the criterion task (Daneman & Carpenter, 1980).

The present studies suggest that when predicting cognitive constructs, a broad set of tasks should be adopted to evaluate the variation in the relationships. First of all, the contribution of external and internal distraction to the relationship between WM performance and cognitive skills differ across criterion tasks, being higher and more consistent in WM-fluid intelligence relationship and absent in WM-GPA relationship. Additionally, the commonality analyses in Study II indicated, the WM span tasks and the criterion tasks used in the present study are assessing highly overlapping cognitive constructs. According to the positive manifold, the cognitive constructs tend to correlate with each other and are difficult to extract (Kovacs & Conway, 2016). Baddeley and Hitch (1974, pp. 86) already pointed out that understanding of the role of WM in cognitive tasks "must proceed hand-in-hand with an understanding of the tasks themselves". Following their suggestion, in differential studies, careful analysis of not only WM tasks, but also the outcomes that are predicted is needed in understanding the environmental or task factors that makes the WM task more predictive of cognitive skills.

Not only associations between measured WM capacity and performance in cognitive tasks, but also the memory scores in the complex span tasks did differ in the function distraction. However, only internal distraction contributed to the memory scores (Study II) as predicted by the controlled attention hypothesis (Engle, 2002; Kane et al., 2001). Memory scores did not differ in the function of external distraction, in this case environmental speech- or non-speech noise (Study I). This discrepancy might result from the differences in the designs of the present studies. The internal distraction was induced in a within-subject design by carefully controlling the demand of the task by varying the amount of information to process while keeping the presentation rate constant (for a similar approach, see Barrouillet et al., 2004). The environmental distraction, in turn, was induced by non-controlled, natural classroom context, in which participants were exposed to different distraction (between subjects). The effects of irrelevant sound on memory (Salamé & Baddeley, 1982) and cognitive performance such as reading (Vasilev, Kirkby, & Angele, 2018) are widely documented, and it is possible, that the setting in the present study was not sufficient in producing differences between the memory scores.

4.2 DOMAIN-GENERAL AND DOMAIN-SPECIFIC TRAINING OF COUNTING SKILLS

Given the strong, robust relationship between WM and scholastic skills shown in wide previous literature and supported by Studies I and II of present thesis, it would be expected to see transfer effects from WM training to scholastic skills. However, Studies III and IV evaluated the evidence and found no effects of WM training on early mathematical skills when training the distinct domains of WM or WM and counting more generally. The result is in line with a majority of earlier studies, supporting the view of limited benefits of WM training (Melby-Lervåg & Hulme, 2013) and demonstrating the advantage of domain-specific interventions in improving emergent numerical skills. The results are in accordance with the predominant view that generalisation of training across different cognitive tasks, that is far transfer, is not commonly observed. Various explanations have been proposed for the lack of transfer. It has been suggested that if WM training would lead to increase in cognitive skills, it should address the precise processes that are shared with cognitive skills, that is, the capacity to hold items held in primary memory, the attentional inhibitory process or binding (von Bastian & Oberauer, 2013). Thus, assuming that the training regimen was sufficient, it seems that either 1) the WM training paradigms did not engage with these skills, 2) that WM just is not malleable or 3) some other aspects mediate the relationship between WM and scholastic skills, such as content knowledge.

Although recently the effects of WM training have raised considerable interest, the issue has a long history. Already Woodworth and Thorndike (1901) suggested that skill acquisition is based on domain-specific knowledge which does not generalise to other domains. Later, it was suggested that the more specialized the skill is, the less there is overlap between skills (Ericsson & Kintsch, 1995), and the more difficult the transfer will be. Most recently, Gathercole and her colleagues (2019) suggested that training-induced transfer is only possible, when the training involves learning of novel cognitive routines that can be applied to new tasks. Accordingly, in the present thesis, WM training, which did not have domain-specific overlap with early numerical skills, did not improve those skills. In contrast, when counting skills were addressed by training, the transfer was promoted, which could be interpreted to result from the fact that counting activities share domain-specific knowledge with early numerical skills.

Some researchers have proposed that the lack of transfer reflects the fact that WM and cognitive skills are not causally related (Harrison et al., 2013; Richey, Phillips, Schunn, & Schneider, 2014). This is one plausible hypothesis. First, there could be other factors affecting both skills. For example, it has been suggested that performance on verbal WM tasks merely reflect variations in language skills and that could affect both WM and comprehension (MacDonald & Christiansen, 2002; see also Klem et al., 2015; Melby-Lervåg et al., 2012). Second, there might be other factors affecting the relationship. For example, the persons with superior cognitive abilities could engage more in academic learning activities in a younger age and thus show higher scholastic achievement (see Sala & Gobet, 2017b). Third, the causal relationship could be other way around. This would mean that the academic learning would improve WM (Finch, 2019). As an example, previous studies demonstrate that the age of the appearance of child's first number words predict the later memory for digits (Libertus, Marschik, & Einspieler, 2014).

However, despite the lack of training-effects of WM training in Study III, Study IV showed that training of the domain of the outcome, in this case numeracy, was effective. In interpreting of the results of Study IV, the role of domain-specific content knowledge is a plausible hypothesis in explaining the transfer. In line with the results of present thesis, Miller and Robertson (2011) demonstrated that cognitive training programs, wchich implemented arithmetic games improved the participants' ability to perform simple calculations. Furthermore, in older students, established interventions for academic learning are elaboration of studied material and repeated testing, which are both based on the activating or enriching the LTM representations instead of affecting the domain-general core capacity of WM (Dunlosky et al., 2013; Roediger & Pyc, 2012). Thus, based on the results of Studies III and IV, it is plausible to suggest that establishing and strengthening the domain-specific information in LTM is beneficial in order to enhance scholastic skills, such as numeracy.

4.3 DOMAIN-GENERAL AND DOMAIN-SPECIFIC WM IN SCHOLASTIC SKILLS

Domain-general views of WM, such as executive attention view of WM, suggest that general cognitive capacity, not the task-specific knowledge, explains the individual differences in WM and scholastic skills. The domain-specific views of WM in turn see WM as a workspace for integrating skills, knowledge, and procedures which are needed in learning particular task, such as in specific mathematical task or in comprehending a text on particular subject (Ericsson & Delaney, 1999). The partial contribution of external and internal distraction on WM and its relationship with fluid intelligence and particular scholastic skills support the domain-general view of WM. However, the lack of training-effects after WM training and instead being present after domain-specific training, suggests that domain-specific aspects are needed in explaining the individual differences in scholastic skills. Thus, results of the four studies of the present thesis support the view of domain general attentional capacity, but also that the domain-specific aspects of WM are in important role in understanding scholastic learning.

These conclusions are supported by recent research, which have explored the relative importance of domain-specific and domain-general factors in scholastic skills (e.g., Fuchs et al., 2010; Geary, 2011; Passolunghi & Lanfranchi, 2012). For example, in the context of reading texts, Kaakinen and her colleagues (2003) demonstrated that with familiar texts, the high-span subjects were able to make use of their prior knowledge more efficiently compared to the low-span subjects. Furthermore, when the texts were unfamiliar, the subjects with high-span were better at allocating their attentional resources to relevant information than subjects with low-span. Similarly, the importance of both domain-general and domain-specific factors in complex tasks has been shown in mathematics: children's mathematical skills are predicted not only by domain-general capacity such as WM capacity, but also by skills specific to mathematics such as symbolic number processing (Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013). Thus, the results suggest that both the ability to use prior knowledge, as suggested Ericsson & Kintsch (1995), and the ability to efficiently control attentional resources (Engle et al., 1999) are common factors in explaining the relationship between reading and WM capacity. Furthermore, recent findings suggest that domain-general and domain-specific skills may have reciprocal effects on each other during development (Peng et al., 2016; Schmitt, Geldhof, Purpura, Duncan, & McClelland, 2017).

4.4 LIMITATIONS

There are some limitations in the present study that suggest ideas for future research. First of all, in Study I, the occurrences of the environmental noises during the course of the assessment were not controlled. The children completed the tasks at their own pace and the occurences of environmental noise was documented and later classified. Thus, the the exact temporal locations of the noises in relation to the tasks and the results reveal the role of noise only at a general level. Experimental paradigms are needed if the aim is to show which cognitive phases and processes are disrupted by which noise at what point in time, specifically in regard to the encoding, maintaining and rehearsal of information (Elliott, 2002). Furthermore, since the study is correlational in nature, it is not possible to tell whether the environmental noise was indeed causing the effects. For example, it is possible that some school classes are overall noisier than others and that the participants were exposed to more noise also when completing paper-and-pencil tests on fluid intelligence, arithmetic, and reading. However, despite these limitations, the results obtained in the natural classroom setting reveal that WM span tasks can be sufficiently assessed in school environments in this age group.

In Studies III and IV in turn, one central limitation is, that it is possible, that there were not enough training sessions and thus the effects would be absent due to the limited amount of training. Both studies had relatively light training regimens lasting for four to five weeks including training twice a week. Earlier studies suggest, that variations in the amount and intensity of training might explain the contrasing results of training studies (Morrison & Chein, 2011). Still, the Study IV showed that unlike WM training, the counting training was beneficial with this small amount of training sessions. Thus, the present results indicate that the number of sessions was sufficient in order to have some change in cognitive abilities, in this case numeracy. This improvement is important in kindergarten-context and has practical value when pedagogical and neuropsychological interventions are developed.

In Studies III and IV there were many methodological choices for avoiding the pitfalls of previous studies, for example the use of active control groups, the adoptation of broad set of outcome measures and the Bayesian analysis of data. Also, total samples were quite large in these studies (160 children participating in two studies). However, the number of participants in each training group were quite small (N = 14–23), which lowers the statistical power of both studies. The research in the field of WM training is highly active and more research continues to be conducted in the field. In future meta-analyses, systematic reviews and umbrella reviews (overviews of systematic reviews in the field; Aromataris et al., 2017) may give more justification for evaluating the benefits of interventions in children and adults.

4.5 WORKING MEMORY BEYOND LABORATORY

In the present thesis, both WM assessment and WM interventions were taken out of the laboratory into more natural setting of school class and kindergarten group. For example, Study I investigated whether the WM can be assessed with mobile tablets and whether the assessment is reliable in the natural environment which is accompanied with ambient speech and non-speech noise. It was shown that the WM scores obtained from the group-based assessment in a classroom with noise were even better predictors of scholastic skills than the scores obtained with no noise. Furthermore, Study IV successfully applied counting intervention performed in small groups. It integrated plays and games suitable for preschool and kindergarten settings. In the WM training studies training is typically conducted in controlled setting, requiring specific amount of training in specific settings. However, context and conditions rarely are similar in real life than in research setting and that has important consequences in interpreting the results (Logie, 2018). For example, people might not find time or motivation to proceed along with the given training schedule in their home environments.

The discrepancy in the research on WM, as in many other fields, is that WM is used in everyday environments, but traditionally studied in a laboratory

(Logie, 2018). Conducting studies in natural environments has potential in increasing the ecological validity of the obtained information. Moreover, it helps in applying the acquired information in school learning, since it decreases the disparity between the research setting and the school setting. Thus, it is important to acquire knowledge from real life and natural environments. According to present research, bringing the research of the role of WM in scholastic skills from the laboratory to the natural environments offers accurate information of the role of WM in scholastic skills.

4.6 FUTURE DIRECTIONS OF RESEARCH ON WM AND SCHOLASTIC SKILLS

The four studies of the present thesis show that performance in WM tasks have some practical value in considering how children and adolescent learn in real life and in explaining the variability in how children and adolescent manage in the school. First of all, the results suggest that assessing children's WM functioning is possible in a real life setting such as in a classroom. Attentional demands posed by external or internal distraction does not lower, but can improve the tasks ability to predict scholastic skills. Thus, for reliably recognising children's and adolescents' WM deficits, relatively cost-effective methods of group-based tablet assessments are currently available. In addition, the results suggest that to support the children struggling with scholastic skills, at least in emergent mathematical skills, practicing those skills at kindergarten and school is relevant.

One recently presented interesting hypothesis for explaining the effects of training on the performance on other cognitive tasks suggests that to be effective, the training should require formulation of new cognitive routines that can be applied to new tasks (Gathercole et al., 2019). This framework assumes that repeated exposure to the WM span task, like any other cognitive task, leads to learning of coordinated sequences of processes, which eventually become autonomous cognitive routines. These cognitive routines can be applied to tasks that share common structure with the trained task. In the context of complex span tasks, it would mean that although dissimilar in the content of processing component and memory component, the tasks share the coordination of the process of altering sequence of those components. Extensive repeating the task allows for those routines to become autonomous. These cognitive routines could be, according to this hypothesis, transferred to tasks that share similar structure. However, transfer would be unlikely on tasks that consist of broad and altering set of cognitive operations, such as academic skills.

In future studies, the hypothesis of cognitive routines could be applied not only to training, but also to WM assessment. Following this idea, it could be hypothesised that WM span tasks are good predictors of complex cognition because we are generally unexperienced with such tasks, and there are no cognitive routines to apply in these tasks (Gathercole et al., 2019). Thus, the question would be: would practising such tasks affect their predicitive utility. If the forming of applicable cognitive routines is crucial in explaining training effects, the forming of new routines by practicing WM tasks in intensive training (near-transfer) should contribute to the correlations between the measured WM span and scholastic skills. Consquently, further studies would benefit from applying the idea of cognitive routines in understanding the factors making WM span tasks such a strong predictor of cognitive tasks. In sum, although WM training has not gained much success in producing far-transfer as in Studies III and IV of present thesis, other studies have shown near-transfer (Sala et al., 2019). Further studies could benefit from combining the differential and experimental methods (Underwood, 1975) in investigating what the consequences of training of WM task for the relationship between WM and cognitive skills.

4.7 CONCLUSIONS

The question of the role of WM in complex cognition, such as in scholastic skills, has intrigued researchers since the introducing of the concept of WM in the seminal paper of Baddeley and Hitch (1974). By combining the information from designs studying group and individual differences, the four studies of present thesis demonstrate that distraction contribute to the relationships between WM and scholastic skills and that school related skills, such as emergent mathematical ability is supported by task specific training rather than domain-general WM training. These results suggest that while attentional load in complex span tasks contributes markedly to the individual differences in the measured WM capacity by restricting the mental workspace, acquired LTM representations also contribute to cognitive skills by enhancing the ability to use WM in scholastic learning. Together, the results of the present thesis question the view that the role of WM restricts in serving purely as general cognitive capacity underlying scholastic skills. In contrast, the results suggest that WM should be seen as a workspace for integrating skills, knowledge, and procedures which are needed in learning particular tasks. Thus, in scholastic tasks, such as in specific mathematical task or in comprehending a text on particular subject WM should be seen as a system with restricted capacity to integrate information. The individual differences not only in the WM capacity, but also in the previous experiences explain this ability.

APPENDIX

Availability of the materials of the present thesis

Study I

WM span tasks: https://github.com/kiistala/working-memory-span-tasks-for-tablet

Analysis scripts for the data analysis presented in the original paper:

https://osf.io/w5qb6/

Study II

No materials, data or scirpts currently openly available.

Study III

Analysis scripts of the BF analysis presented in the present thesis:

https://osf.io/hqyk5/

Published paper:

http://www.diva-portal.org/smash/get/diva2:955121/FULLTEXT01.pdf#page=160

Study IV

Analysis scripts of the BF analysis presented in the present thesis:

https://osf.io/hqyk5/

REFERENCES

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General, 131*, 567–589. doi:10.1037/0096-3445.131.4.567.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30–60. doi:10.1037/0033-2909.131.1.30.

Ahonen T., & Kiuru N. (2013-2017). The Stairway Study–From Primary School to Secondary School [Unpublished data set]. University of Jyväskylä, Finland.

Alloway, T. P. (2006). How does working memory work in the classroom? *Educational Research and Reviews*, *1*, 134-139.

Alloway, T. P. (2007). *Automated Working: Memory Assessment: Manual*. London, UK: Pearson Assessment.

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, 106, 20–29. doi:10.1016/j.jecp.2009.11.003.

Alloway, T. P., Gathercole, S. E., Adams, A. M., Willis, C., Eaglen, R., & Lamont, E. (2005). Working memory and phonological awareness as predictors of progress towards early learning goals at school entry. *British Journal of Developmental Psychology*, *23*, 417-426.

Aromataris E., Fernandez R., Godfrey C., Holly C., Khalil H., & Tungpunkom P. (2017). Chapter 10: Umbrella Reviews. In E. Aromataris, Z. Munn (Eds.), *Joanna Briggs Institute Reviewer's Manual*. The Joanna Briggs Institute. Retrieved from https://reviewersmanual.joannabriggs.org/

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic Bulletin & Review*, *22*, 366-377.

Aunola, K., & Räsänen, P. (2007). The 3-minutes Basic Arithmetic Test. Unpublished test material. University of Jyväskylä, Finland.

Baddeley, A. (1986). Working memory. New York: Oxford University Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory?. *Trends in Cognitive Sciences*, *4*, 417-423.

Baddeley, A. D. (2002). Is working memory still working? *European Psychologist*, *7*, 85-97.

Baddeley, A. (2007). *Working Memory, Thought, and Action*. New York: Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of Learning and Motivation,* 8, 47-89. doi:10.1016/S0079-7421(08)60452-1.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, *128*, 612-637.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General*, *133*, 83-100. doi:10.1037/0096-3445.133.1.83.

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General*, *132*, 71-92. doi:10.1037/0096-3445.132.1.71.

Bunting, M. (2006). The role of processing difficulty in the predictive utility of working memory span. *Psychonomic Bulletin and Review, 13*, 998–1004. doi:10.3758/BF03213915.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*, 365-376. doi:10.1038/nrn3475.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386-404.

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163-183.

Conway, A. R. A., Jarrold, C., & Kane, M. J., Miyake, A., & Towse, J. N. (2007). *Variation in Working Memory*. New York: Oxford University Press.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*, 769-786.

Conway, A. R. A., & Kovacs, K. (2013). Individual differences in intelligence and working memory: A review of latent variable models. *Psychology of Learning and Motivation*, 58, 233–270. doi:10.1016/B978-0-12-407237-4.00007-4.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*, 197-223.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*, 1158-1170.

Cowan, N., Morey, C.C., AuBuchon, A.M., Zwilling, C.E., & Gilchrist, A.L. (2010). Seven year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science, 13*, 120-133.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450-466.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*, 13–21. doi:10.1016/j.Intell.2006.02.001.

De Simoni, C., & von Bastian, C. C. (2018). Working memory updating and binding training: Bayesian evidence supporting the absence of transfer. *Journal of Experimental Psychology: General*, *147*, 829-859. doi:10.1037/xge0000453.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5:* 781.

Dougherty, M. R., Hamovitz, T., & Tidwell, J. W. (2016). Reevaluating the effectiveness of n-back training on transfer through the Bayesian lens: Support for the null. *Psychonomic Bulletin & Review*, *23*, 306-316.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology.

Psychological Science in the Public Interest, 14, 4–58. doi:10.1177/1529100612453266.

Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition, 30*, 478-487. doi:10.3758/BF03194948.

Elliott, E. M., & Briganti, A. M. (2012). Investigating the role of attentional resources in the irrelevant speech effect. *Acta Psychologica, 140*, 64–74. doi:10.1016/j.actpsy.2012.02.009.

Elliott, E. M., Hughes, R. W., Briganti, A., Joseph, T. N., Marsh, J. E., & Macken, B. (2016). Distraction in verbal short-term memory: Insights from developmental differences. *Journal of Memory and Language, 88*, 39–50. doi:10.1016/j.jml.2015.12.008.

Engle, R. W. (2001). What is working memory capacity. *The Nature of Remembering: Essays in Honor of Robert G. Crowder*, 297-314. Washington, US: American Psychological Association. doi:10.1037/10394-016.

Engle, R. W. (2002). Working memory capacity as executive attention. Current *Directions in Psychological Science, 11*, 19-23.

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*, 190-193.

Engle, R. W., & Kane, M. J. (2003). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of Learning and Motivation, 145–199*. doi:10.1016/s0079-7421(03)44005-x.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in WMC and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge, England: Cambridge University Press.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*, 309.

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 257-297). New York: Cambridge University Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245. doi:10.1037/0033-295X.102.2.211

Fenesi, B., Sana, F., Kim, J. A., & Shore, D. I. (2015). Reconceptualizing working memory in educational research. *Educational Psychology Review*, *27*, 333-351.

Finch, J. E. (2019). Do Schools Promote Executive Functions? Differential Working Memory Growth Across School-Year and Summer Months. *AERA Open, 5*, 2332858419848443.

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development*, *81*, 1520-1533.

Gathercole, S. E., Brown, L., & Pickering, S. J. (2003). Working memory assessments at school entry as longitudinal predictors of National Curriculum attainment levels. *Educational and Child Psychology, 20*, 109-122.

Gathercole, S. E., Dunning, D. L., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of Memory and Language*, *105*, 19-42.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology*, *70*, 177-194.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, 40, 177-190. doi:10.1037/0012-1649.40.2.177.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*, 1-16.

Gathercole, S. E., Woolgar, F., Kievit, R. A., Astle, D., Manly, T., Holmes, J., & CALM Team. (2016). How common are WM deficits in children with difficulties in reading and mathematics? *Journal of Applied Research in Memory and Cognition, 5*, 384-394.

Geary, D. C. (2011). Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. *Journal of Developmental and Behavioral Pediatrics*, *32*, 250-263. doi:10.1097/DBP.0b013e318209edef.

Geers, A. E., Pisoni, D. B., & Brenner, C. (2013). Complex working memory span in cochlear implanted and normal hearing teenagers. *Otology & Neurotology*, *34*, 396-401. doi:10.1097/MAO.0b013e318277a0cb.

Gignac, G. E. (2014). Fluid intelligence shares closer to 60% of its variance with working memory capacity and is a better indicator of general intelligence. *Intelligence*, 47, 122–133.doi:10.1016/j.intell.2014.09.004.

Guye, S., & von Bastian, C. C. (2017). Working memory training in older adults: Bayesian evidence supporting the absence of transfer. *Psychology and Aging*, *32*, 732-746.

Harrison, T. L., Shipstead, Z., Hicks, K. L., Hambrick, D. Z., Redick, T. S., & Engle, R. W. (2013). Working memory training may increase working memory capacity but not fluid intelligence. *Psychological Science, 24*, 2409-2419.

Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: individual differences in visual attention control. *Journal of Experimental Psychology: General, 136*, 217-240. 10.1037/0096-3445.136.2.217.

Hirvonen, K. (2012). Onko laskutaito laskussa? Matematiikan oppimistulokset peruskoulun päättövaiheessa 2011. Koulutuksen seurantaraportti 12:4. Helsinki: Opetushallitus.

Holopainen, L., Kairaluoma, L., Nevala, J., Ahonen, T., & Aro, M. (2004). Lukivaikeuksien seulontamenetelmä nuorille ja aikuisille [Dyslexia screening test for youth and adults]. Jyväskylä, Finland: Niilo Mäki Instituutti.

Hygge, S. (2003). Classroom experiments on the effects of different noise sources and sound levels on long-term recall and recognition in children. *Applied Cognitive Psychology, 17*, 895-914. doi:10.1002/acp.926.

Imbo, I., & Vandierendonck, A. (2008). Effects of problem size, operation, and working-memory span on simple-arithmetic strategies: differences between children and adults?. *Psychological Research*, *72*, 331-346.

Jeffreys, H. (1961). Theory of probability (3rd ed.). New York: Oxford University Press.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2003). How prior knowledge, WMC, and relevance of information affect eye fixations in expository text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 447-457. doi:10.1037/0278-7393.29.3.447.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*,*130*, 169-183.

Kane, M. J., Conway, A. R. A, Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds) *Variation in Working Memory* (pp. 21-48). New York: Oxford University Press.

Kievit, R. A., Lindenberger, U., Goodyer, I. M., Jones, P. B., Fonagy, P., Bullmore, E. T., ... & Dolan, R. J. (2017). Mutualistic coupling between vocabulary and reasoning supports cognitive development during late adolescence and early adulthood. *Psychological Science*, *28*, 1419-1431.

Klein, K., & Fiss, W. H. (1999). The reliability and stability of the Turner and Engle working memory task. *Behavior Research Methods, Instruments, & Computers*, *31*, 429-432.

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science*, *18*, 146-154.

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences*, *14*, 317-324.

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, *24*, 781-791.

Kovacs, K., & Conway, A. R. A. (2016). *Process Overlap Theory: A Unified Account of the General Factor of Intelligence. Psychological Inquiry, 27,* 151–177. doi:10.1080/1047840x.2016.1153946.

Krumm, S., Ziegler, M., & Buehner, M. (2008). Reasoning and working memory as predictors of school grades. *Learning and Individual Differences, 18*, 248–257. doi:10.1016/j.lindif.2007.08.002.

Kudo, M. F., Lussier, C. M., & Swanson, H. L. (2015). Reading disabilities in children: A selective meta-analysis of the cognitive literature. *Research in Developmental Disabilities*, *40*, 51-62.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?!. *Intelligence*, *14*, 389-433.

Kyttälä, M., Kanerva, K., Munter, I., & Björn, P. M. (2019). Working memory resources in children: stability and relation to subsequent academic skills. *Educational Psychology*, 1-20. doi:10.1080/01443410.2018.1562046.

Lampit, A., Hallock, H., & Valenzuela, M. (2014). Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. *PLoS medicine*, *11*, e1001756.

Landerl, K., Wimmer, H., & Moser, E. (1997). Salzburger Lese-und rechtschreibtest (SLRT). Verfahren zur Differentialdiagnose von Störungen des Lesens und Schreibens für die, 1.

Lee M. D. & Wagenmakers E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge, UK: Cambridge University Press.

Lépine, R., Barrouillet, P., & Camos, V. (2005). What makes working memory spans so predictive of high-level cognition? *Psychonomic Bulletin & Review, 12*, 165–170. doi:10.3758/BF03196363.

Libertus, M. E., Marschik, P. B., & Einspieler, C. (2014). Number word use in toddlerhood is associated with number recall performance at seven years of age. *PloS one*, *9*, e98573. doi:10.1371/journal.pone.0098573.

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*, 861-883.

Logie, R. (2018). Human cognition: Common principles and individual variation. *Journal of Applied Research in Memory and Cognition*, *7*, 471-486.

Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, *22*, 395-410.

Lu, L., Weber, H. S., Spinath, F. M., & Shi, J. (2011). Predicting school achievement from cognitive and non-cognitive variables in a Chinese sample of elementary school children. *Intelligence, 39*, 130-140.

Lustig, C., Hasher, L., & Zacks, R. T. (2007). Inhibitory deficit theory: Recent developments in a “new view”. *Inhibition in Cognition, 17*, 145-162.

MacDonald, M. C., & Christiansen, M. H. (2002). Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, 109, 35-54. doi:10.1037//0033-295X.109.1.35.

Maehler, C., & Schuchardt, K. (2009). Working memory functioning in children with learning disabilities: does intelligence make a difference?. *Journal of Intellectual Disability Research, 53*, 3-10.

Magimairaj, B. M., & Montgomery, J. W. (2012). Children’s verbal working memory: Role of processing complexity in predicting spoken sentence comprehension. *Journal of Speech, Language, and Hearing Research, 55*, 669–682. doi:10.1044/1092-4388(2011/11-0111).

McCabe, D. P. (2010). The influence of complex working memory span task administration methods on prediction of higher level cognition and metacognitive control of response times. *Memory & Cognition, 38*, 868–882. doi:10.3758/MC.38.7.868.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37*, 1–10. doi:10.1016/j.intell.2008.08.004.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, *49*, 270-292. doi:10.1037/a0028228.

Melby-Lervåg, M., Lervåg, A., Lyster, S. A. H., Klem, M., Hagtvet, B., & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children’s vocabulary development. *Psychological Science*, *23*, 1092-1098.

Menghini, D., Finzi, A., Carlesimo, G. A., & Vicari, S. (2011). Working memory impairment in children with developmental dyslexia: is it just a phonological deficity?. *Developmental Neuropsychology*, *36*, 199-213.

Miller, D. J., & Robertson, D. P. (2011). Response to Logie and Della Sala: Brain training in schools, where is the evidence?. *British Journal of Educational Technology*, *42*, E101-E102.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*, 621-640. doi:10.1037//0096-3445.130.4.621.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control.* Cambridge, UK: University Press.

Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationships among working memory, secondary memory, and fluid intelligence. *Psychological science, 19*, 1071-1077.

Morey, R. D., Rouder, J. N., Jamil, T., & Morey, M. R. D. (2015). Package 'bayesfactor'. Retrieved from *http://cran/r-projectorg/web/packages/BayesFactor/BayesFactor*

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin & Review, 18*, 46-60.

Nevala, J., Kairaluoma, L., Ahonen, T., Aro, M., & Holopainen, L. (2006). Lukemis- ja kirjoittamistaitojen yksilötestistö nuorille ja aikuisille [Individual test material for assessing dyslexia in youth and in adult age]. Jyväskylä, Finland: Niilo Mäki Instituutti.

Nevo, E., & Bar-Kochva, I. (2015). The relations between early working memory abilities and later developing reading skills: a longitudinal study from kindergarten to fifth grade. *Mind, Brain, and Education*, *9*, 154-163, doi:10.1111/mbe.12084.

Nimon, K., & Reio, T. (2011). Regression commonality analysis: A technique for quantitative theory building. *Human Resource Development Review, 10*, 329–340. doi:10.1177/1534484311411077.

Oberauer, K., Farrell, S., Jarrold, C., Pasiecznik, K., & Greaves, M. (2012). Interference between maintenance and processing in working memory: The effect of item–distractor similarity in complex span. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 665-685. doi:10.1037/a0026337.

Palladino, P., & Ferrari, M. (2008). Phonological sensitivity and memory in children with a foreign language learning difficulty. *Memory*, *16*, 604-625. doi:10.1080/09658210802083072.

Passolunghi, M. C., & Lanfranchi, S. (2012). Domain-specific and domain-general precursors of mathematical achievement: A longitudinal study from kindergarten to first grade. *British Journal of Educational Psychology*, *82*, 42-63.

Peng, P., & Fuchs, D. (2016). A meta-analysis of working memory deficits in children with learning difficulties: Is there a difference between verbal domain and numerical domain?. *Journal of Learning Disabilities*, *49*, 3-20. doi:10.1177/0022219414521667.

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., ... Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin*, 144, 48-76. doi:10.1037/bul0000124.

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology*, *108*, 455-473.

Peng, P., Wang, T., Wang, C., & Lin, X. (2019). A meta-analysis on the relation between fluid intelligence and reading/mathematics: Effects of tasks, age, and social economics status. *Psychological Bulletin*, *145*, 189-236. doi:10.1037/bul0000182.

Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children (WMTB-C)*. Psychological Corporation.

Preßler, A. L., Krajewski, K., & Hasselhorn, M. (2013). Working memory capacity in preschool children contributes to the acquisition of school relevant precursor skills. *Learning and Individual Differences*, *23*, 138-144.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*, 110-122.

Raven, J., Court, J., & Raven, J. (1995). *Coloured Progressive Matrices*. Oxford, England: Oxford Psychologists Press.

Raven, J., Raven, J. E., & Court, J. H. (1998). *Standard Progressive Matrices*. Oxford, England: Oxford Psychologists Press.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*, 164-171.

Richey, J. E., Phillips, J. S., Schunn, C. D., & Schneider, W. (2014). Is the link from working memory to analogy causal? No analogy improvements following working memory training gains. *PloS One, 9,* e106616.

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence, 35*, 83–92. doi:10.1016/j.intell.2006.05.004.

Roediger, H. L., Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, 1, 242–248. doi:10.1016/j.jarmac.2012.09.002.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356-374.

Räsänen, P., Salminen, J., Wilson, A. J., Aunio, P., & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development, 24*, 450–472. doi:10.1016/j.cogdev.2009.09.003.

Sala, G., Aksayli, N. D., Tatlidil, K. S., Tatsumi, T., Gondo, Y., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. *Collabra: Psychology, 5*, 18. doi:10.1525/collabra.203.

Sala, G., & Gobet, F. (2017a). Working memory training in typically developing children: A meta-analysis of the available evidence. *Developmental Psychology, 53*, 671-685.

Sala, G., & Gobet, F. (2017b). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science, 26*, 515-520.

Sala, G., & Gobet, F. (2018). Cognitive training does not enhance general cognition. *Trends in Cognitive Sciences,* 1856, 1-12.

Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21*, 150–164. doi:10.1016/S0022-5371(82)90521-7.

Salonen, P., Lepola, J., Vauras, M., Rauhanummi, T., Lehtinen, E., & Kinnunen, R. (1994). Diagnostiset testit 3. Motivaatio, metakognitio ja matematiikka. Turku, Finland: Oppimistutkimuksen keskus.

Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number–space mappings: What underlies mathematics achievement?. *Journal of Experimental Child Psychology*, *114*, 418-431.

Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R., & McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *Journal of Educational Psychology*, *109*, 1120-1140. doi:10.1037/edu0000193.

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist*, *50*, 138-166.

Seigneuric, A., & Ehrlich, M. F. (2005). Contribution of working memory capacity to children's reading comprehension: A longitudinal investigation. *Reading and Writing*, *18*, 617-656. doi:10.1007/s11145-005-2038-0.

Service, E. (1992). Phonology, working memory, and foreign language learning. *The Quarterly Journal of Experimental Psychology, 45*, 21–50.

Shelton, J. T., Elliott, E. M., Matthews, R. A., Hill, B. D., & Gouvier, W. M. (2010). The relationships of working memory, secondary memory, and general fluid intelligence: working memory is special. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 813-820. doi:10.1037/a0019046.

Shipstead, Z., Harrison, T. L., & Engle, R. W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science, 11*, 771-799.

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: A multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review, 24,* 1077-1096.

St. Clair-Thompson, H. L. (2007). The effects of cognitive demand upon relationships between working memory and cognitive skills. *The Quarterly Journal of Experimental Psychology, 60*, 1378–1388. doi:10.1080/17470210601025505.

Sullivan, J. R., Osman, H., & Schafer, E. C. (2015). The effect of noise on the relationship between auditory working memory and comprehension in school-age children. *Journal of Speech, Language, and Hearing Research, 58*, 1043-1051.

Süss, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability – and a little bit more. *Intelligence, 30*, 261–288. doi:10.1016/S0160-2896(01)00100-3.

Sörqvist, P. (2010). The role of working memory capacity in auditory distraction: A review. *Noise & Health, 12*, 217–224. doi:10.4103/1463-1741.70500.

Sörqvist, P., Stenfelt, S., & Rönnberg, J. (2012). Working memory capacity and visual–verbal cognitive load modulate auditory–sensory gating in the brainstem: Toward a unified view of attention. *Journal of Cognitive Neuroscience, 24*, 2147–2154. doi:10.1162/jocn_a_00275.

Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology*, *103*, 821-837. doi:10.1037/a0025114.

Swanson, H. L., Howard, C. B., & Saez, L. (2006). Do different components of working memory underlie different subgroups of reading disabilities? *Journal of Learning Disabilities, 39*, 252–269.

Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology, 100*, 343.

Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology*, *79*, 294-321.

Takeuchi, H., Taki, Y., & Kawashima, R. (2010). Effects of working memory training on cognitive functions and neural systems. *Reviews in the Neurosciences*, *21*, 427-450.

Tremblay, S., Nicholls, A. P., Alford, D., & Jones, D. M. (2000). The irrelevant sound effect: Does speech play a special role? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1750-1754. doi: 10.1037/0278-7393.26.6.1750.

Towse, J. N., & Hitch, G. J. (2007). Variation in working memory due to normal development. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds) *Variation in Working Memory* (pp. 109-133). NY: Oxford University Press.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent?. *Journal of Memory and Language*, *28*, 127-154.

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist*, *30*, 128-134.

Unsworth, N. (2010). Interference control, working memory capacity, and cognitive abilities: A latent variable analysis. *Intelligence, 38*, 255-267.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133*, 1038-1066.

Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: individual differences in voluntary saccade control. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30(6), 1302.

Van Luit, J. E. H., Van de Rijt, B. A. M. & Aunio, P. (2006) Early Numeracy Test, Finnish edition [Lukukäsitetesti]. Helsinki, Finland: Psykologien kustannus.

Vasilev, M. R., Kirkby, J. A., & Angele, B. (2018). Auditory distraction during reading: A Bayesian meta-analysis of a continuing controversy. *Perspectives on Psychological Science*, *13*, 567-597. doi: 10.1177/1745691617747398.

Vock, M., Preckel, F., & Holling, H. (2011). Mental abilities and school achievement: A test of a mediation hypothesis. *Intelligence, 39*, 357–369. doi:10.1016/j.intell.2011.06.006.

von Bastian, C. C., & Oberauer, K. (2013). Distinct transfer effects of training different facets of working memory capacity. *Journal of Memory and Language*, *69*, 36-58.

von Bastian, C. C., & Oberauer, K. (2014). Effects and mechanisms of working memory training: a review. *Psychological Research*, *78*, 803-820.

Wechsler, D. (2010). *Wechsler Intelligence Scale for Children IV.* Helsinki, Finland: Psykologien Kustannus Oy.

Wetzel, N., & Schröger, E. (2007). Modulation of involuntary attention by the duration of novel and pitch deviant sounds in children and adolescents. *Biological Psychology, 75*, 24–31. doi:10.1016/j.biopsycho.2006.10.006.

Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions. (I). *Psychological Review, 8*, 247–261. doi:10.1037/h0074898.